DETECTION VS SELECTION: INTEGRATION OF GENETIC, EPIGENETIC AND ENVIRONMENTAL CUES IN FLUCTUATING ENVIRONMENTS

John M. McNamara[1*], Sasha R. X. Dall[2], Peter Hammerstein[3] & Olof Leimar[4]

1. School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK; john.mcnamara@bristol.ac.uk

2. Centre for Ecology & Conservation, Biosciences, College of Life & Environmental Sciences, University of Exeter, Penryn Campus, Tremough, Penryn TR10 9EZ, UK; sashadall@iname.com

3. Department of Biology, Humbolt University, Invalidenstr. 43, 10115 Berlin, Germany; p.hammerstein@biologie.hu-berlin.de

4. Department of Zoology, Stockholm University, SE-106 91 Stockholm, Sweden; olof.leimar@zoologi.su.se

Author contributions: Inspired by conversations with OL, JMM formulated the model (with modifications suggested by the other authors), performed the analyses and obtained the main results. All authors contributed to the literature search and writing the ms.

Short running title: Integration of genetic and environmental cues.

Keywords: Transgenerational effects, plasticity, bet hedging, maternal effects, habitat tracking, adaptive development, reaction norm.

Submitted as a Letter to Ecology Letters.

Number of words in abstract: 150.

Number of words in main text: 4945.

Number of references: 31.

Number of display items: 5 figures + 1 table.

*Corresponding author: Tel +44 117 9287986, Fax +44 117 9287999, email john.mcnamara@bristol.ac.uk.

ABSTRACT

There are many inputs during development that influence an organism's fit to current or upcoming environments. These include genetic effects, transgenerational epigenetic influences, environmental cues and developmental noise, which are rarely investigated in the same formal framework. We study an analytically-tractable evolutionary model, in which cues are integrated to determine mature phenotypes in fluctuating environments. Environmental cues received during development and by the mother as an adult act as detection-based (individually observed) cues. The mother's phenotype and a quantitative genetic effect act as selection-based cues (they correlate with environmental states after selection). We specify when such cues are complementary and tend to be used together, and when using the most informative cue will predominate. Thus, we extend recent analyses of the evolutionary implications of subsets of these effects by providing a general diagnosis of the conditions under which detection and selection-based influences on development are likely to evolve and coexist.

INTRODUCTION

Organisms are sensitive to a variety of inputs during development, often producing phenotypes that are suited to current or upcoming environments (West-Eberhard, 2003). Adaptive phenotypic plasticity and transgenerational effects are among the well-studied examples. In spatially varying environments, genetic variation contributes to local adaptation (e.g., Levene, 1953; Seger & Brockmann, 1987; Kawecki & Ebert, 2004) and allele frequencies will vary spatially. An individual's genotype will therefore statistically contain information about local environmental conditions and thus can be regarded as a genetic cue that can be combined and integrated with environmental and transgenerational cues adaptively during development (Lively, 1986; Sultan & Spencer, 2002; Leimar *et al.*, 2006; Leimar & McNamara, 2015; Dall *et al.*, 2015). In temporally fluctuating environments on the other hand, it is traditional to consider only environmental cues and, sometimes, transgenerational cues as being the developmental influences that fit phenotypes to current conditions. Random phenotype determination (diversified bet hedging) is another important adaptation to unpredictable environments (Seger & Brockmann, 1987). Both evolutionary modeling (Lachmann & Jablonka, 1996) and empirical observation (Bergland *et al.*, 2014; Cogni *et al.*, 2015) indicate that genetic variation plays a role in fitting phenotypes to temporal environmental variation, provided that the time scale of variation is longer than the generation time of the organism. However, such genetic effects are rarely investigated alongside trangenerational effects, bet hedging and adaptive plasticity in the same formal framework. Here we study the relative importance and interaction of all of these influences on phenotype determination in temporally varying environments, using an evolutionary model that, to a great extent, can be worked out analytically.

Amongst trans-generational epigenetic effects, Shea *et al.* (2011) make the distinction between *detection-based* and *selection-based* effects. The former are con-

cerned with the influence of cues about environmental conditions that are directly observed and are passed down the generations. Indeed, current environmental influences on adaptive development can also be thought of as detection-based effects (equivalent to information by "instruction": Jablonka & Lamb (2005)). In contrast, selection based effects do not require direct observations by individuals. They occur when there is transmission (with reasonable fidelity) of an epigenetic marker down successive generations, where the marker affects the phenotype, and so is under selection, and as a result of past selection current individuals tend to adaptively match their environment. Such selection-based effects (via heritable genetic variation) also form the basis of the "genes-as-cues" analysis of Leimar *et al.* (2006) and Leimar & McNamara (2015) in spatially heterogeneous environments. Here, for the first time, we explore the relative value of using the full range of potential detection-based and selection-based cues during development in temporally varying environments in the same model. In environments without spatial structure it is only favorable for the parental generation to pass information to offspring when environments are autocorrelated (so that knowledge of the environment in one year is predictive of the environment in the folllowing year) (Shea *et al.*, 2011; Kuijper *et al.*, 2014; English *et al.*, 2015; Uller *et al.*, 2015; Kuijper & Hoyle, 2015). In this case the maternal phenotype or cue genes can act as selection-based sources of information, and this is the situation we study.

In our model there are two environmental sources of information (Figure 1) that act as direct detection-based cues; during development each individual receives a cue of the current environment ("juvenile cue") that can affect the adult phenotype; in addition each individual receives a further environmental cue as an adult that can be passed to offspring. Both cues are subject to noise and so are not perfectly informative. There are also two selection-based cues; the phenotype of the mother and a quantitative genetic effect, present in the offspring, that, in our model acts as a cue to the offspring, but can also be seen as a breeding value for the trait

in question. The phenotype of an individual can depend on its juvenile cue, its mother's phenotype, the mother's adult cue and its inherited quantitative genetic cue genes, as well as developmental noise, the level of which is under selection (Figure 1). As in Rivoire & Leibler (2014), our analysis involves two different timescales; the environment, the distribution of maternal phenotypes and that of the genetic cue genes all vary from generation to generation, whereas the developmental system that integrates these cues is passed on to offspring without error. We seek the developmental system that maximises the long-term growth rate in the number of individuals that employ this means of phenotype determination.

It has been previously suggested that the use of a genetic cue determines its correlation with the environment and hence value as a cue (Leimar, 2009); a use it or lose it principle. For the first time we give an explicit demonstration of this principle. However, our main focus is on the interaction of the various cues, and how this interaction depends on the rate of environmental change and the accuracy of cues and information transmission. Some previous models (Rivoire & Leibler, 2014; English *et al.*, 2015; Leimar & McNamara, 2015) have considered combinations of cues, but our model, which considers a specific purely temporarily varying environment, allows an analytic expression for fitness and is, we believe, particularly suited to exposing the logic of cue integration. Unlike Rivoire & Leibler (2014) we allow separate inheritance channels so as to have a clear separation and analysis of the effects of selection based versus detection based cues, which are otherwise entangled. Although the synergy between detection and selection based cues has been previously proposed (e.g., Kuijper & Hoyle, 2015), we give the first clear demonstration of the positive synergy between environmental cues and the maternal phenotype; the combination of these cues results in much higher fitness than when only one of these cues is used. In contrast, environmental cues and cue genes do not synergise in the same way and incorporating both does not always result in higher fitness. Thus, unlike recent models that analyse the evolutionary implications of subsets of

cues (e.g., Leimar *et al.*, 2006; Shea *et al.*, 2011; Kuijper *et al.*, 2014; Leimar & McNamara, 2015; Kuijper & Hoyle, 2015; Kuijper & Johnstone, 2016) our analysis provides a general diagnosis of the conditions under which detection and selection-based influences on development are likely to evolve and/or coexist.

METHODS

We assume an asexual population with discrete, non-overlapping generations. There are two genetically determined elements. One is a quantitative effect that acts as a genetic cue to the developmental system. The other is the cue integration system itself. This system determines how the genetic cue, maternal phenotype, two types of environmental cues and noise jointly influence development and hence determine the adult phenotype. We allow the quantitative genetic trait to evolve for a given cue integration system, finding the fitness of the cue integration system. We then find the cue integration system with the greatest fitness. Model details are similar to that of Rivoire & Leibler (2014). Both models allow the influence of detection based cues to be inherited (a form of Lamarkism) but in Rivoire & Leibler (2014) the mother passes a single quantity on to her offspring. This quantity is a linear combination of the maternal phenotype, the cue received by the mother as an adult and the quantity passed on to the mother by her mother. In contrast, we allow for the maternal phenotype, the adult maternal cue and the genetic cue genes to be passed on to offspring separately before the offspring combines them to determine its phenotype (Figure 1).

The environment. The environmental state in generation t is $\theta(t)$. The dynamics are given by

$$\theta(t+1) = \lambda\theta(t) + \epsilon_\theta(t). \tag{1}$$

Here $0 < \lambda < 1$ and $\epsilon_\theta(t) \sim N(0, \sigma^2)$ is independent of current and previous environmental states. The stochastic process $\{\theta(t); t = 0, 1, 2, \cdots\}$ is then a stationary Markov process with an equilibrium distribution that is normally distributed with mean 0 and variance

$$\mathrm{Var}(\theta) = \frac{\sigma^2}{1 - \lambda^2}. \tag{2}$$

The parameter λ is the correlation coefficient between environmental states at successive times; i.e.

$$\rho(\theta(t+1), \theta(t)) = \lambda. \tag{3}$$

The genetic cue. The quantitative genetic effect can take any real value. Surviving offspring of a parent with genetic effect value z' have effect value $z = z' + \epsilon_Z$, where $\epsilon_Z \sim N(0, \sigma^2_{mut})$.

Environmental cues. A juvenile in generation t receives two environmental cues that can affect its mature phenotype; it observes the juvenile cue C_J (where $C_J \sim N(\theta, \sigma^2_J)$) and is passed the cue C_A (where $C_A \sim N(\theta(t-1), \sigma^2_A)$) that its mother observed as an adult (Figure 1). These cues are conditionally independent given these environmental states. Within a generation the cues received by different population members are also conditionally independent resulting in a distribution of cues that is centred on the current environmental state. This distribution varies across generations as the environment varies.

Phenotype determination. The adult phenotype of an individual is given by

$$x = \alpha z + \beta_J c_J + \beta_A c_A + \gamma(m + \epsilon_m) + \delta\epsilon_\delta, \tag{4}$$

where z is the value of its genetic effect, c_J is its juvenile environmental cue, c_A is the adult environmental cue observed by her mother, m is the phenotype of the mother, $\epsilon_m \sim N(0, \sigma^2_m)$ is the error in transmission of the maternal phenotype to the offspring and $\epsilon_\delta \sim N(0, 1)$ is a developmental noise term. Here α, β_J, β_A, γ and

δ are non-negative genetically determined parameters that specify the action of the developmental system.

Reproductive success. Reproductive success is a function of the fit of the phenotype to the environment; specifically an individual of phenotype x leaves

$$Ke^{-\frac{1}{2}(x-\theta)^2} \tag{5}$$

surviving offspring when the environmental state is θ. Here K is a positive constant.

Fitness. We evaluate the geometric mean fitness $G(\alpha, \beta_J, \beta_A, \gamma, \delta)$ of the developmental system. Consider a large (essentially infinite) cohort of individuals with this developmental system. Let $X(t)$ be the phenotype of a randomly selected cohort member and $\bar{X}(t)$ the mean phenotype in generation t. We show (Supporting Information, Section SI.1) that if within a generation the joint distribution of $X(t)$ and the quantitative genetic effect are bivariate normal then they remain so in future generations. We thus assume that the distribution of $X(t)$ given $\bar{X}(t) = \bar{x}$ is normal with mean $\bar{x}$ and variance σ_X^2. We also argue (SI.2) that this variance tends to a limiting stationary value, and we assume the cohort has achieved this value. Let $\theta(t) = \theta$ and $\bar{X}(t) = \bar{x}$. Then, since the cohort is large (so that we can average over demographic stochasticity), between generation t and $t+1$ the cohort grows by the factor

$$R(\theta, \bar{x}) = K\mathbb{E}(e^{-\frac{1}{2}(X(t)-\theta(t))^2}|\theta, \bar{x}). \tag{6}$$

Thus, using the fact that the conditional distribution of $X(t)$ is normal we have

$$R(\theta, \bar{x}) = \frac{K}{\sqrt{1+\sigma_X^2}} e^{-\frac{1}{2}\frac{(\bar{x}-\theta)^2}{1+\sigma_X^2}}. \tag{7}$$

The geometric mean fitness of the developmental system is

$$G(\alpha, \beta_J, \beta_A, \gamma, \delta) = e^{\mathbb{E}(\ln R(\theta, \bar{X})}, \tag{8}$$

where the expectation is taken with the respect to the stationary distribution of the vector process $\{(\theta(t), \bar{X}(t)) : t = 0, 1, 2, \ldots\}$. Since $\bar{X} - \theta$ is symmetric about zero, and hence has mean 0 (SI.4) we have $\mathbb{E}((\bar{X} - \theta)^2) = \mathrm{Var}(\bar{X} - \theta)$. Thus

$$G(\alpha, \beta_J, \beta_A, \gamma, \delta) = \frac{K}{\sqrt{1 + \sigma_X^2}} \exp\left[-\frac{1}{2}\frac{\mathrm{Var}(\bar{X} - \theta)}{1 + \sigma_X^2}\right]. \tag{9}$$

Note that fitness depends on the weights $\alpha, \beta_J, \beta_A, \gamma, \delta$ through their influence on both σ_X and $\mathrm{Var}(\bar{X} - \theta)$. We denote the values of these weights that maximise fitness by $\alpha^*, \beta_J^*, \beta_A^*, \gamma^*, \delta^*$.

RESULTS

Diversified bet hedging

Suppose that individuals receive no information on the current environmental state ($\alpha = \beta_J = \beta_A = \gamma = 0$) so that phenotype determination is given by $x = \delta\epsilon_\delta$. Then the best fixed trait value is $x = 0$ since the environment is symmetric about $\theta = 0$. However, always maturing with this phenotype is not a robust strategy when the environmental variance is large, and a strategy that incorporates diversified bet hedging will achieve greater geometric mean fitness (cf. Seger & Brockmann (1987)). Specifically, in SI.5 it is shown that the optimal phenotype determination is given by $x = \delta^*\epsilon_\delta$, where $\delta^* = 0$ for $\mathrm{Var}(\theta) < 1$ and $\delta^* = \sqrt{\mathrm{Var}(\theta) - 1}$ for $\mathrm{Var}(\theta) \geq 1$.

Environmental cue during development

Suppose that juveniles receive a cue during development but no other cue so that $x = \beta_J c_J + \delta\epsilon_\delta$. Since different individuals receive different cues (whose distribution centres on the underlying environmental state), producing a range of phenotypes within a generation, the need to have additional diversified bet hedging is removed and $\delta^* = 0$ (SI.6). Thus we can restrict attention to phenotype determination of the

form $x = \beta_J c_J$.

Two values of β_J have an obvious statistical interpretation. $\beta_J = 1$ corresponds to using the minimum variance unbiased estimator for θ. This estimator has mean θ for all θ but has high variance within a generation. The arithmetic mean (over θ) annual growth in genotype numbers is maximised by setting $\beta_J = \beta_{bayes}$, where $\beta_{bayes} = \mathrm{Var}(\theta)/(\mathrm{Var}(\theta) + \sigma_J^2)$ is the Bayes posterior mean for θ given cue c_J. This method of phenotype determination results in a large discrepancy between the mean phenotype within a generation and θ when $|\theta|$ is large, and consequently has a high variance in annual growth. As Figure 2a illustrates, the optimal value of β_J is a compromise between these two values; i.e. $\beta_{bayes} < \beta_J^* < 1$. [See SI.6 for a proof.]

Environmental cue received by the mother as an adult

If an individual's only cue is that experienced by its mother as an adult (i.e. $x = \beta_A c_A$), the value of this cue depends on the likely change in the environment between the maternal and the current generation. As a result, the optimal weight put on this cue increases with increasing environmental autocorrelation λ (Figure 2b).

When an individual receives both adult maternal and juvenile cues during development ($x = \beta_J c_J + \beta_A c_A$), it can be shown that $\beta_J^* + \lambda\beta_A^* < 1$ (SI.8). Since the juvenile cue is more up-to-date more weight should be placed on it when both cues have the same cue error variance; although as the environmental autocorrelation increases to its maximum value of 1 the weights become equal (Figure 2b). Similar effects of the degree of environmental stability were obtained by English *et al.* (2015) and Leimar & McNamara (2015). As can be seen from Figure 2c, in this example the juvenile cue is more important in terms of fitness than the maternal adult cue when λ is low. Both cues contribute significantly to fitness for high environmental autocorrelation.

Genetic cue

Suppose that the quantitative genetic effect is the only available cue and there is no randomisation, so that phenotype determination is given by $x = \alpha z$. If this cue is ignored ($\alpha = 0$), there is no selection on the genetic effect and its value is uninformative. As α increases the selection pressure on the genetic effect increases resulting in an increased correlation between the effect and the environmental state (Figure 3a), so that the effect acts as a selection-based source of information. In other words, the more notice is taken of the genetic effect the more informative is its value, leading to a feedback in which it should be used more. This feedback is limited; fitness declines for high α (Figure 3b) since too high a value leads to too much variation in the phenotype within a generation (high σ_X^2, cf. equation (9)).

As the environmental autocorrelation increases for given $\mathrm{Var}(\theta)$, so that the environment varies more slowly but has the same variability, selection leads to a higher correlation between the genetic effect and the environmental state (Figure 3a), leading to an increase in fitness (Figure 3b). This is in contrast to the effect of λ for a purely juvenile cue.

Regardless of what combination of cues is available, the fitness of the optimal developmental system does not depend on the mutation rate of the effect genes since an increase in the mutation rate is equivalent to a proportionate decrease in the parameter α; fitness depends on α and σ_{mut} only through the product $\alpha\sigma_{mut}$. (This can be deduced from SI.2 - SI.4.) When there is just the genetic cue it may be optimal to have some randomisation ($\delta^* > 0$). The range of environmental parameters for which randomisation is optimal is explored in Rivoire & Leibler (2014).

Figures 3c,d illustrate optimal phenotype determination when there is both a genetic and a juvenile cue. In this case no additional randomisation is required ($\delta^* = 0$). Figure 3c illustrates the optimal norm of reaction to the juvenile cue for two values of the genetic effect. As can be seen, the slope of the norm of reaction is less and influence of the genetic effect is stronger when the environmental autocorrelation

is higher. Figure 3d shows the amount of phenotypic variation that is attributed to the influence of each cue. (Since fitness depends on the parameter α only through $\alpha\sigma_{mut}$, in presenting results we have shown the breakdown of total variance rather than showing α^*.) For low values of the environmental autocorrelation λ the genetic cue is not used even though this cue would have been used had the juvenile cue not been available, illustrating a certain lack of synergy between these cues. As λ increases the amount of phenotypic variation due to the influence of the genetic cue increases rapidly and that due to juvenile cue falls sharply. Further computation (not shown) reveal that the value of λ below which the genetic cue is ignored increases as the variance in the juvenile cue decreases.

Maternal phenotype as a cue

Since the reproductive success of the mother depends on the fit between her phenotype and the environment, the fact that an individual has been born suggests that her mother's phenotype was close to the environmental state. Thus maternal phenotype can act as a selection-based source of information during development.

When the maternal phenotype is the only developmental cue, there is error-free transmission of information on the maternal phenotype to offspring ($\sigma_m^2 = 0$) and no developmental noise ($\delta = 0$), all phenotypes quickly reduce to $m = 0$ and the maternal phenotype becomes uninformative. Thus in order that the maternal phenotype contains useful statistical information, it is necessary to include transmission error or developmental noise so as to maintain variation within a generation which selection can act on. This can be seen as a timescale issue; if there is no variation the developmental system is committed to existing in a single phenotype, which is then an evolutionary dead end when the environment changes. By incorporting variation the developmental system always ensures that at all future times it is present in some individuals that do well.

As Figure 4a shows, when variation is maintained the correlation between mater-

nal phenotype and the environmental state increases with increasing λ, increasing the value, and hence the weight, put on the maternal phenotype as a cue, and less developmental noise is required (Figure 4b). Transmission error perfectly substitutes for developmental noise, provided the variation generated by our chosen transmission error does not exceed that which is optimal (Figure 4b). In all cases, fitness is a strictly increasing function of λ (Figure 5b). Note that, unlike the model of Kuijper & Johnstone (2016), successive environments are always positively autocorrelated in our model so that we always have $\gamma^* \geq 0$.

As Figure 4b illustrates, we always have $\delta^* > 0$ when $\sigma_m^2 = 0$. In particular, even though $\delta^* = 0$ when the phenotype is determined by $x = \delta\epsilon_\delta$ when $\mathrm{Var}(\theta) \leq 1$ (see above), we have $\delta^* > 0$ when phenotype determination is via $x = \gamma m + \delta\epsilon_\delta$; illustrating the synergy between noise and the influence of the maternal phenotype.

Maternal phenotype and juvenile cue: cross-generational environmental cue integration

We now consider the case where an individual can respond to the environmental cue during development (the juvenile cue) and to the phenotype of her mother. It has previously been suggested that the maternal phenotype may encapsulate previous environmental cues (Townley & Ezard, 2013; Kuijper & Hoyle, 2015). To investigate this effect in this context and to motivate the form of trait determination, let c_0, c_{-1}, c_{-2}, ... be the juvenile cues received by the individual, her mother, her grandmother, and so on. During development it would be clearly advantageous, but not realistic, for an individual to have available all the juvenile cues received by its ancestors. It might nevertheless be reasonable to assume some suitable summary of these past cues is passed on. To explore this idea we note that, in the absence of censoring (due to differential mortality), it is straightforward to show that the Bayes posterior

mean of the current environmental state given $c_0, c_{-1}, c_{-2}, c_{-3}, \cdots$ can be written as

$$\hat{c}_0 = (1-\kappa)[c_0 + \kappa\lambda c_{-1} + (\kappa\lambda)^2 c_{-2} + (\kappa\lambda)^3 c_{-3} + \cdots], \tag{10}$$

where the constant κ is a function of λ, σ and σ_J (cf. Townley & Ezard (2013)). This posterior mean is a sufficient statistic for the current environmental state, and can be written as $\hat{c}_0 = (1-\kappa)c_0 + \kappa\lambda\hat{c}_{-1}$, where $\hat{c}_{-1}$ is the corresponding posterior mean for the mother. Assuming the phenotype determination satisfies $x = \hat{\beta}_J \hat{c}_0$, we can thus write this trait as $x = (1-\kappa)\hat{\beta}_J c_J + \kappa\lambda m$, where, in keeping with previous notation, we now denote the current juvenile cue c_0 by c_J and the phenotype of the mother by m. This analysis shows that if phenotype determination is of the form $x = \beta_J c_J + \gamma m$ then the maternal phenotype provides information in two different ways. As before it provides selection-based information, but now that there is a juvenile cue, it also encapsulates information from previous juvenile cues. This increases the correlation between maternal phenotype and the current environmental state (Figure 4a), and hence increases the value of the maternal phenotype as a cue. Consequently the maternal phenotype should always be used as a cue ($\gamma^* > 0$) when both are available (SI.9). This is in contrast to the combination of maternal adult cue and maternal phenotype, when it can be the case that $\gamma^* = 0$ (Uller *et al.*, 2015).

As the environmental autocorrelation increases the maternal phenotype becomes a more valuable cue both because the past selective environment has been more stable and because past juvenile cues are more relevant to current conditions. Thus under optimal phenotype determination more weight is given to the maternal phenotype as a cue and less to the current juvenile cue (Ezard *et al.*, 2014; Uller *et al.*, 2015), although relative weights depend on cue error variances and the fidelity in transmission of the maternal phenotype (Figure 4c).

Comparison of genetic and maternal cues

Figure 5a illustrate how cues perform in combination. When the maternal phenotype is the only cue fitness is very similar to that when the genetics effect is the only cue (the maternal cue is slightly superior as it is transmitted to offspring without error here, whereas the genetic effect mutates), so that the two cues are essentially interchangeable. Furthermore, very little is gained by allowing both cues at the same time. However, the situation is completely different when there is a juvenile environmental cue; the synergy between this cue and the maternal phenotype results in significantly higher fitness than the combination of juvenile cue and genetic cue, which have no synergy. Furthermore, the genetic effect is not used when this third cue is available in this setting. Adding the genetic effect to the other two cues is rarely advantageous, although its inclusion increases fitness slightly when there is developmental noise and λ is very close to 1 (not illustrated).

Adding noise to the transmission of the maternal phenotype reduces the advantage of the maternal cue and juvenile cue combination (Figure 5b), but this combination remains superior to that of the genetic and environmental cue even when there is considerable noise unless the environmental autocorrelation is close to $\lambda = 1$.

The combination of juvenile and adult maternal environmental cues is inferior to the combination of juvenile and maternal phenotype (Figure 5a) unless there is significant error in transmission of her phenotype (Figure 5b), since the maternal phenotype encapsulate information on earlier environments. Furthermore adding the adult maternal cue to the juvenile and maternal phenotype only produced a small increase in fitness (Figure 5a).

Ecological conditions			Detection based		Selection based	
Env. auto-correlation	Env. cues	Phenotype inheritance	Juvenile cue	Adult cue	Maternal pheno-type	Genes
low	any	any	✓	✗	✗	✗
high	accurate	very inaccurate	✓✓	✓✓	✗	✓
high	inaccurate	very inaccurate	✓	✓	✗	✓✓
high	inaccurate	accurate	✓	✓	✗✗ ✓✓	✓✓ ✗✗
high	accurate	accurate	✓✓	✓✓	✓✓	✗

Table 1: The combination of cues that is predicted under various combinations of factors (the degree of environmental autocorrelation, the accuracy of environmental cues and the accuracy with which the mother's phenotype can be passed to offspring). A single tick denotes significant selection pressure to use a cue, a cross denotes very weak selection pressure, double ticks or crosses denote very strong or extremely weak pressure, respectively. Under the fourth condition there are two alternative best methods of phenotype determination; rely heavily on the maternal phenotype or rely on genes (but not both). Note that although the maternal phenotype is categorised as a selection-based cue, it can incorportate detection-based information (see text).

DISCUSSION

We allow the development of an individual to be affected by four cues. Two are directly observed environmental cues; a juvenile cue that the individual experienced during development and a cue experienced by her mother as an adult and passed to the individual. Two are selection based cues; a quantitative genetic effect and the phenotype of the mother. The three cues passed on from the mother use separate inheritance channels (Figure 1) so as to give a clear separation of the effects of selection-based versus detection-based cues. Our main focus is on the interaction of the various cues, and how this depends on environmental variance and autocorrelation, the accuracy of environmental cues and the accuracy of transmission of the maternal phenotype. We give the first clear demonstration of the positive synergy between environmental cues and the maternal phenotype and lack of synergy between environmental and cue genes. In addition we show that the juvenile cue can

act as a randomisation device, analyse the feedback between use of a genetic cue and its value as a cue, and highlight issues of timescale. Table 1 summarises the relative influences of the cues that are predicted by our model.

When the environmental autocorrelation is low the mature phenotype mainly depends on the juvenile cue since the maternal adult cue is out of date (Figure 2a) and the selection-based cues are poorly correlated with the current environmental state (see e.g. Figure 3a). There is a strong dependence on the juvenile cue if it is accurate, but even an inaccurate cue acts as a source of phenotype diversification and removes the need to bet hedge via developmental noise when the environmental variance is high. In our model cues received by different population members are uncorrelated given the environment. For example, if the environmental state represented mean food availability, the actual amount found by different individuals might be centred on this mean but vary in an uncorrelated way due to good and bad luck when foraging. However, any cue, even if inaccurate, which gave a spread of estimates of the environmental state, could potentially obviate the need to have truly randomised phenotype determination; although noise in gene expression will inevitably introduce some randomisation in development (Eldar & Elowitz, 2010).

When the environmental autocorrelation is high and there is high error in the transmission of the information on the maternal phenotype to offspring, the quantitative genetic effect is always used as a cue, although the relative weight put on this cue depends on both the strength of the autocorrelation and the accuracy of the two environmental cues.

The strength of selection on cue genes increases with their influence in development; a "use it or lose it" principle that we demonstrate for the first time (Figure 3a). This result relates to the finding of Kawecki (2000) that the effect of a modifier changes the selection on structural genes. Analogous feedback also occurs in models of phenotype determination in spatially heterogeneous environments. For example, if population members are natally philopatric then they tend to be born

in local habitats to which they are already adapted, so that it can be optimal to be natally philopatric, ignoring developmental cues that have significant probability of error (McNamara & Dall, 2011). If, however, population members took notice of such cues they would disperse more and might not be particularly adapted to their birth habitat. It would then be better to take notice of developmental cues; i.e. not be natally philopatric. The presence of feedbacks raises the possibility that there may be more than one local fitness optima (Dall *et al.*, 2015), although the fitness landscapes appears unimodal in the cases illustrated in Figure 3.

When the environment is highly autocorrelated and there is high fidelity in the inheritance of information on the maternal phenotype, both genetic and maternal-phenotype selection-based cues act in a similar manner, and are alternative means of phenotype determination, when detection-based cues are inaccurate (Figure 5a). However, these interact in very different ways with the detection-based environmental cues when the latter are accurate, since the maternal phenotype acts as a summary of previous detection-based cues (a sort of phenotypic memory sensu Kuijper & Johnstone (2016)). Consequently, the combination of environmental cue and maternal phenotype achieves greater fitness than the combination of environmental cue and genetic cue, provided that the maternal phenotype can be accurately passed on to offspring (Figure 5a). In both our model and that of Rivoire & Leibler (2014) a juvenile cue influences the adult phenotype, which in turn is passed on to offspring. This is essentially a form of Lamarckism (by which we mean the inheritance of detection-based cues). From our analysis, this model of transmission seems to be a very efficient way of integrating information, but real organisms might not have mechanisms that can achieve it with high accuracy (except for cultural inheritance), so Lamarckian effects could be limited by a noisy transmission of the maternal phenotype.

In contrast to the model of Rivoire & Leibler (2014) our model has several channels of transmission from parent to offspring (Figure 1), for instance separate

channels for quantitative genetic effects and adult cues. This often corresponds to biological reality, perhaps as a consequence of evolution latching on to different feasible implementations of transgenerational effects. For instance, a transfer of a substance from mother to offspring might be a mechanism that more readily evolves than an integration of adult cues into the hereditary material. Well studied cases of such mechanisms include the "egg foam factor" that plays a part in the determination of the gregarious morph of desert locusts (Miller *et al.*, 2008), and alpha-Tocopherol (a vitamin E) inducing rotifer morphs by being transmitted to offspring (Gilbert, 2016).

We have not explicitly investigated the role that the strength of selection might play, but previous work has shown that selection-based cues become more informative as selection increases in strength (Leimar *et al.*, 2006; Kuijper & Hoyle, 2015), although in contrast, Uller *et al.* (2015) (equation 2.21 and below) find that inheritance of the maternal phenotype (through incomplete resetting of an epigenetic mark) is favoured when selection is weak. In our model we take cue or transmission accuracy as a given parameter. Future work might consider the evolution of channel accuracy. This issue would be expecially important when extending our analysis to social transmission of information and in the more complex case of the transmission of multivariate maternal effects (Townley & Ezard, 2013; Kuijper *et al.*, 2014; Chevin & Lande, 2015). Our model also does not take into account environmental changes during the lifetime of an organism (see, e.g. Nettle *et al.* (2013)). An obvious extension would be to incorporate both changes between and within generations within the same model. In such a setting information passed across generations would act as a Bayesian prior that would then be updated during the lifetime (Stamps & Frankenhuis, 2016).

ACKNOWLEDGMENTS

This work was supported by a Leverhulme Trust International Network Grant to the four authors and by a grant from the Swedish Research Council (621-2010-5437) to O.L.

REFERENCES

Bergland, A. O., Behrman, E. L., O'Brien, K. R., Schmidt, P. S. & Petrov, D. A. (2014). Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in drosophila. *PLOS Genetics*, 10, doi: 10.1371/journal.pgen.1004775.

Chevin, L.-M. & Lande, R. (2015). Evolution of environmental cues for phenotypic plasticity. *Evolution*, 69, 2767–2775.

Cogni, R., Kuczynski, K., Lavington, E., Koury, S., Behrman, L., Brien, K. R. O., Schmidt, P. S., Eanes, W. F., Behrman, E. L., O'Brien, K. R., Schmidt, P. S. & Eanes, W. F. (2015). Variation in drosophila melanogaster central metabolic genes appears driven by natural selection both within and between populations. *Proceedings of the Royal Society B: Biological Sciences*, 282, 10.1098/rspb.2014.2688.

Dall, S. R., McNamara, J. M. & Leimar, O. (2015). Genes as cues: Phenotypic integration of genetic and epigenetic information from a darwinian perspective. *Trends in Ecology & Evolution*, 30, 327–333.

Eldar, A. & Elowitz, M. (2010). Functional roles for noise in genetic circuits. *Nature*, 467, 167–173.

English, S., Pen, I., Shea, N. & Uller, T. (2015). The information value of non-genetic inheritance in plants and animals. *Plos ONE*, 10, 10.1371/journal.pone.0116996.

Ezard, T. H. G., Prizak, R. & Hoyle, R. B. (2014). The fitness costs of adaptation via phenotypic plasticity and maternal effects. *Functional Ecology*, 128, 693–701.

Gilbert, J. J. (2016). Non-genetic polymorphisms in rotifers: environmental and endogenous controls, development, and features for predictable or unpredictable environments. *Biological Reviews*, doi: 10.1111/brv.12264.

Jablonka, E. & Lamb, M. J. (2005). *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral and Symbolic Variation in the History of Life*. MIT Press.

Kawecki, T. J. (2000). The evolution of genetic canalization under fluctuating selection. *Evolution*, 54, 1–12.

Kawecki, T. J. & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, 7, 1225–1241.

Kuijper, B. & Hoyle, R. B. (2015). When to rely on maternal effects and when on phenotypic plasticity? *Evolution*, 69, 950–968.

Kuijper, B. & Johnstone, R. A. (2016). Parental effects and the evolution of phenotypic memory. *Journal of Evolutionary Biology*, 29, 265–276.

Kuijper, B., Johnstone, R. A. & Townley, S. (2014). The evolution of multivariate maternal effects. *PLoS Comput. Biol.*, 10, e1003550.

Lachmann, M. & Jablonka, E. (1996). The inheritance of phenotypes: An adaptation to fluctuating environments. *Journal of Theoretical Biology*, 181, 1–9.

Leimar, O. (2009). Environmental and genetic cues in the evolution of phenotypic polymorphism. *Evolutionary Ecology*, 23, 125–135.

Leimar, O., Hammerstein, P. & Van Dooren, T. J. M. (2006). A new perspective on developmental plasticity and the principles of adaptive morph determination. *American Naturalist*, 167, 367–376.

Leimar, O. & McNamara, J. M. (2015). The evolution of transgenerational integration of information in heterogeneous environments. *American Naturalist*, 185, E55–E69.

Levene, H. (1953). Genetic equilibrium when more than one ecological niche is available. *American Naturalist*, 87, 331–333.

Lively, C. M. (1986). Canalization versus developmental conversion in a spatially variable environment. *American Naturalist*, 128, 561–572.

McNamara, J. M. & Dall, S. R. X. (2011). The evolution of unconditional strategies via the 'multiplier effect'. *Ecology Letters*, 14, 237–243.

Miller, G. A., Islam, M. S., Claridge, T. D. W., Dodgson, T. & Simpson, S. J. (2008). Swarm formation in the desert locust schistocerca gregaria: isolation and nmr analysis of the primary maternal gregarizing agent. *J. Experimental Biology*, 211, 370–376.

Nettle, D., Frankenhuis, W. & Rickard, I. . (2013). The evolution of predictive adaptive responses in human life history. *Proc R Soc B*, 280, http://dx.doi.org/10.1098/rspb.2013.1343.

Rivoire, O. & Leibler, S. (2014). A model for the generation and transmission of variations in evolution. *Proceedings of the National Academy of Sciences*, 111, E1940–E1949.

Seger, J. & Brockmann, H. J. (1987). What is bet-hedging? *Oxford Surveys in Evolutionary Biology*, 4, 182–211.

Shea, N., Pen, I. & Uller, T. (2011). Three epigenetic information channels and their different roles in evolution. *Journal of Evolutionary Biology*, 24, 1178–1187.

Stamps, J. & Frankenhuis, W. (2016). Bayesian models of development. *Trends Ecol. Evol.*, 31, 260–268.

Sultan, S. E. & Spencer, H. G. (2002). Metapopulation structure favors plasticity over local adaptation. *American Naturalist*, 160, 271–283.

Townley, S. & Ezard, T. H. G. (2013). A g matrix analogue to capture the cumulative effects of nongenetic inheritance. *J Evolutionary Biology*, 26, 370–376.

Uller, T., English, S. & Pen, I. (2015). When is incomplete epigenetic resetting in germ cells favoured by natural selection? *Proc. R. Soc. B*, 282, doi: 10.1098/rspb.2015.0682.

West-Eberhard, M. J. (2003). *Developmental Plasticity and Evolution.* Oxford University Press.

FIGURE CAPTIONS

Figure 1. Phenotype determination of an individual in generation t. In generation $t-1$ the mother receives a juvenile environmental cue during development and then matures, at which time her phenotype is set. Her reproductive success is a function of this phenotype and the current environmental state. She also receives a further environmental cue as an adult. This cue doe not affect her phenotype, which is already set, but is passed on to any offspring in generation t, along with the information about her phenotype and her mutated cue genes. These three cues, together with an environmental cue received as a juvenile, determine the phenotype of the offspring. There are thus two detection-based cues; the adult maternal environmental cue and the juvenile environmental cue, and two selection-based cues; the maternal phenotype and cue genes. Note that if phenotypes are influenced by environmental cues, the mother's phenotype as a cue will combine elements of detection and selection (see text).

Figure 2. Individuals may receive one or both of two environmental cues; a juvenile cue during development and a cue passed on from the mother that the mother received as an adult. (a) Effect of the environmental variance when individuals receive only a juvenile cue. Solid curve: the optimal juvenile cue weight β_J^* for three value of the juvenile cue error variance (top curve $\sigma_J^2 = 0.5$, middle curve $\sigma_J^2 = 2.5$, bottom curve $\sigma_J^2 = 10$). Dashed curve below the corresponding solid curve: the value β_{bayes} such that $\beta_{bayes} c_J$ is the Bayes posterior mean for θ given the cue c_J. [Note that β_J^* is the same for all combinations of the values of σ^2 and λ that result in the same value of $\mathrm{Var}(\theta)$; this result can be derived from the formulae in SI.4]. (b) Effect of the environmental autocorrelation on optimal cue weights. Top two curves: individuals receive just one of the cues. Bottom two curves: individuals receive both cues. (β_J^* solid curve, β_A^* dashed curve.) (c) Effect of the environmental autocorrelation on fitness of the optimal developmental system when only the adult cue to the mother is available (bottom curve), only the juvenile cue is available (middle curve) and both cues are available (top curve). In (b) and (c), cue error variances $\sigma_J^2 = \sigma_A^2 = 2.5$, and as λ increases σ is decreased so that the environmental variance if held fixed at the value $\mathrm{Var}(\theta) = 2.5$.

Figure 3. Quantitative genetic effect as a cue. (a) Correlation between the genetic effect and the environmental state (taken across individuals and generations) when there are no other cues. Solid curve $\lambda = 0.95$, dashed curve $\lambda = 0.85$. Mutation variance $\sigma_{mut}^2 = 1.0$. (Correlations derived from the formulae in SI.4). (b) Fitness of the developmental system for the cases considered in (a). (c) Optimal norms of reaction to the juvenile cue ($x = \alpha^* z + \beta_J^* c_J$) for two values of the genetic effect (blue, genetic effect $z = 0$; red, genetic effect $z = 1$), shown for two values of the environmental autocorrelation (solid curves, $\lambda = 0.95$; dashed curves, $\lambda = 0.85$). (d) Breakdown of the total phenotypic variance (both within and across generations) under optimal phenotype determination. Solid curve: variation ex-

plained by response to the juvenile cue ($(\beta_J^*)^2\sigma_J^2$), dashed curve: variation explained by response to the genetic effect $(\alpha^*)^2\mathrm{Var}(Z)$), dotted curve: variation explained by the interaction between these cues ($2\alpha^*\beta_J^*\mathrm{Cov}(Z, C_J)$). [Here Z and C_J are the genetic effect value and juvenile cue value, respectively, of a randomly selected population member in a random generation.] In all cases σ is chosen so that $\mathrm{Var}(\theta) = 2.5$.

Figure 4. The maternal phenotype as a cue. (a) Correlation between the maternal phenotype and the current environmental state under optimal phenotype determination. Solid curve: when there is no other cue but developmental noise is allowed (so that the phenotype is determined as $x = \gamma^* m + \delta^* \epsilon_\delta$). Dashed curves: when in addition there is also a juvenile cue during development (top dashed curve $\sigma_J^2 = 1.0$, lower dashed curve $\sigma_J^2 = 2.5$. Transmission of maternal phenotype information is error free ($\sigma_m^2 = 0$). (b) Optimal phenotype determination when maternal phenotype is the only cue and there is developmental noise. Solid curve: the weight given to the maternal phenotype γ^*. Dashed curve: the amount of randomisation δ^*. In each case the upper (blue) curve corresponds to transmission of the maternal phenotype without error ($\sigma_m^2 = 0$) and the lower (red) curve to $\sigma_m^2 = 0.5$. (c) Optimal phenotype determination when the maternal phenotype is a cue and there is a juvenile cue. Dashed curves show weights (β_J^*) given to the juvenile cue and solid curves show weights given to the maternal phenotype. Cases illustrated are: (i) $\sigma_J^2 = 1.0$, $\sigma_m^2 = 1.5$, (ii) $\sigma_J^2 = 1.0$, $\sigma_m^2 = 0$, (iii) $\sigma_J^2 = 2.5$, $\sigma_m^2 = 0$. In all figures, as λ increases σ is decreased so that the environmental variance if held fixed at the value $\mathrm{Var}(\theta) = 2.5$.

Figure 5. The fitness under optimal phenotype determination for various combinations of cues. (a) Dotted lines are top: juvenile cue + adult maternal cue, bottom: juvenile cue alone. Other curves are (from top to bottom): maternal phenotype + juvenile cue + adult maternal cue, maternal phenotype + juvenile cue,

genetic cue + juvenile cue, maternal phenotype + genetic cue, maternal phenotype alone, genetic cue alone. $\sigma_m^2 = 0$ throughout. Random phenotype determination is allowed although $\delta^* = 0$ except for the maternal phenotype alone case. (b) Dashed curve is for the combination of the genetic cue and juvenile cue. Other curves are all for the combination of maternal cue and juvenile cue, with the error of transmission of the maternal cue (σ_m^2) increasing from top to bottom (cases shown, $\sigma_m^2 = 0, 0.1, 0.25, 1.0, 2.5$). In both figures, as λ increases σ is adjusted so that $\mathrm{Var}(\theta) = 2.5$. Juvenile cue error variance $\sigma_J^2 = 2.5$. Adult maternal cue error variance $\sigma_A = 2.5$.

Figure 1

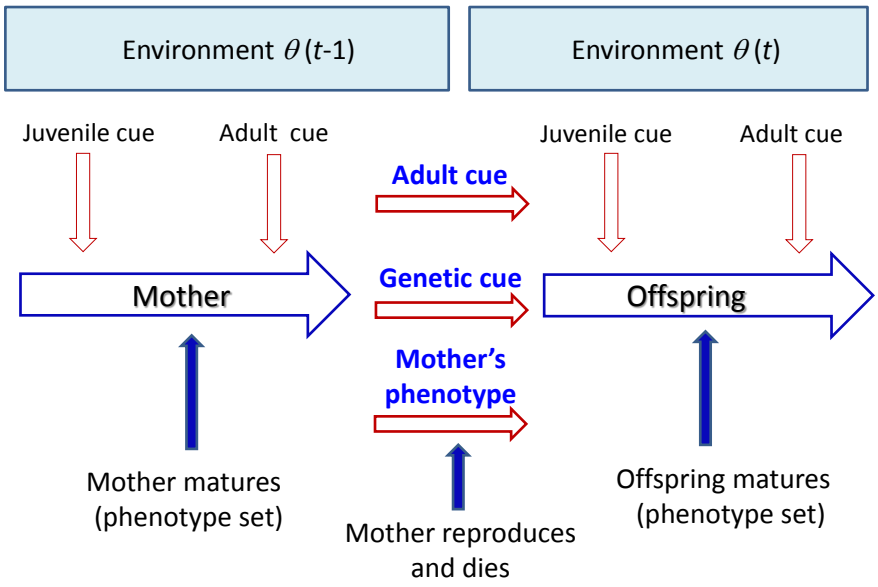

Environment $\theta(t\text{-}1)$
Environment $\theta(t)$
Juvenile cue
Adult cue
Juvenile cue
Adult cue
Adult cue
Genetic cue
Mother's phenotype
Mother
Offspring
Mother matures (phenotype set)
Mother reproduces and dies
Offspring matures (phenotype set)

Figure 2a

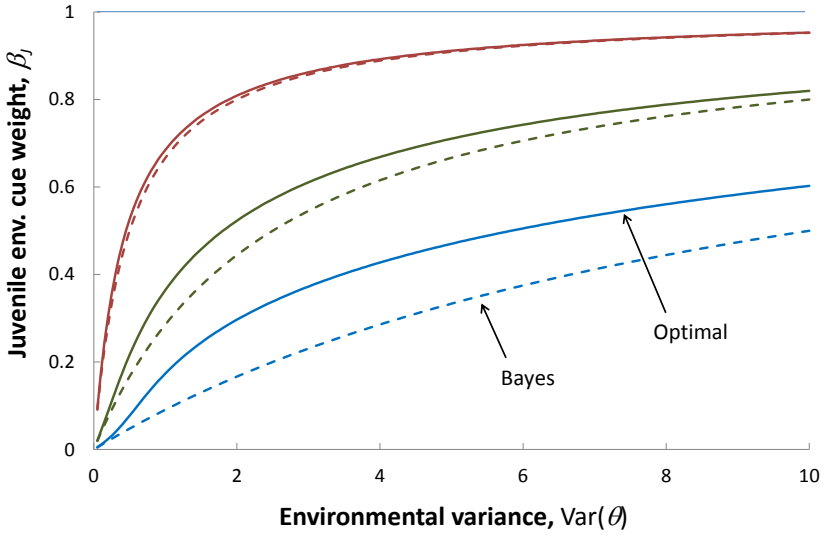


Figure 2b

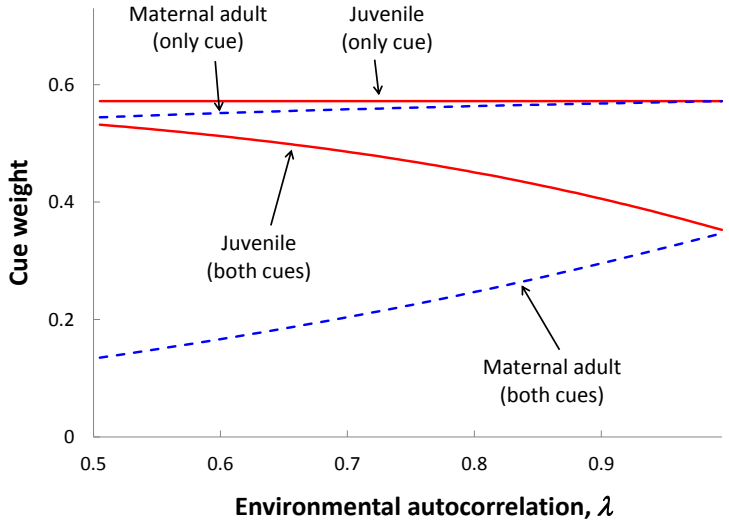


Figure 2c

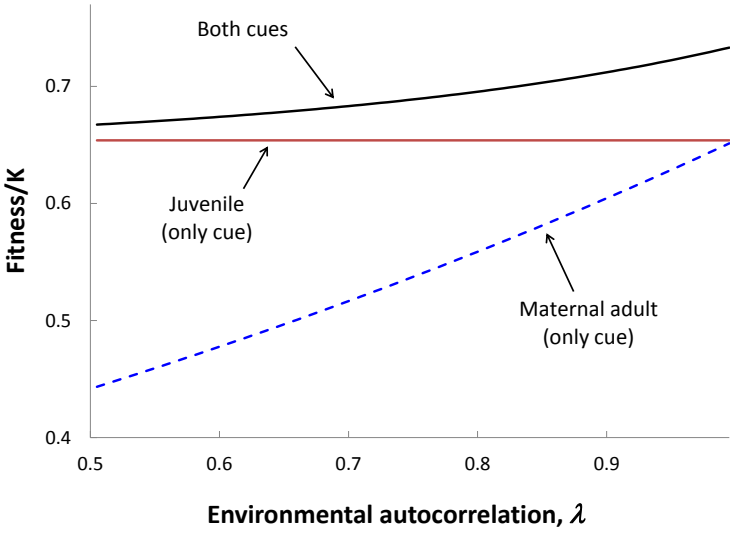

Both cues
Juvenile
(only cue)
Maternal adult
(only cue)
Fitness/K
0.7
0.6
0.5
0.4
0.5
0.6
0.7
0.8
0.9
Environmental autocorrelation, λ

Figure 3a

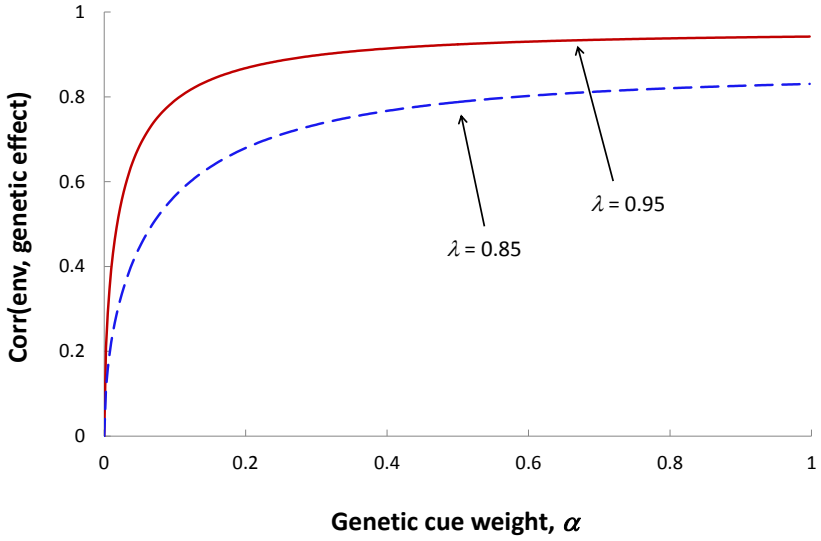


Figure 3b

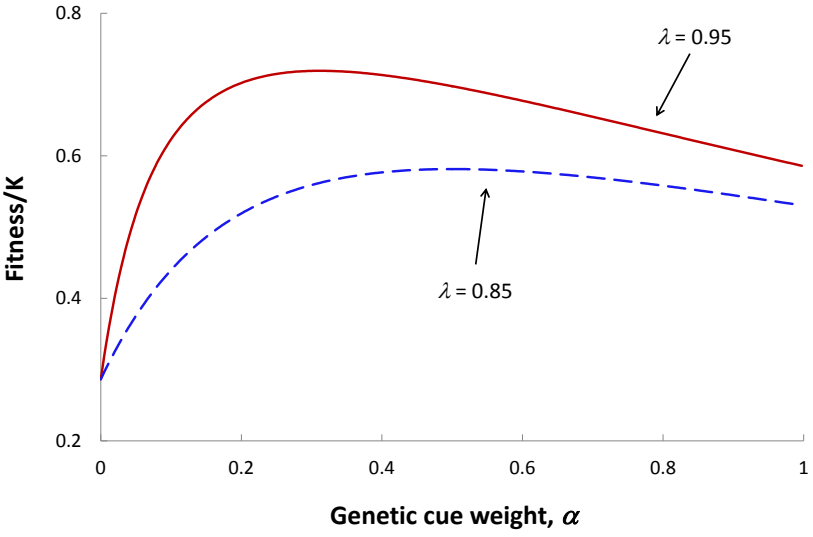


Figure 3c

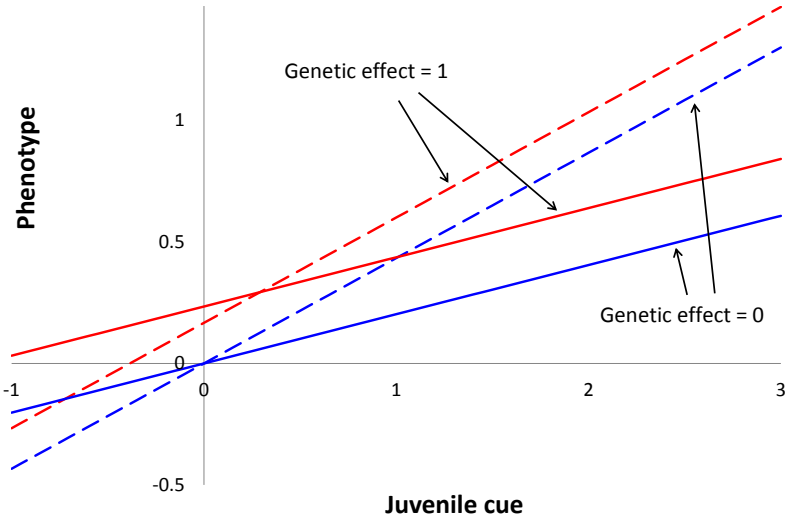


Figure 3d

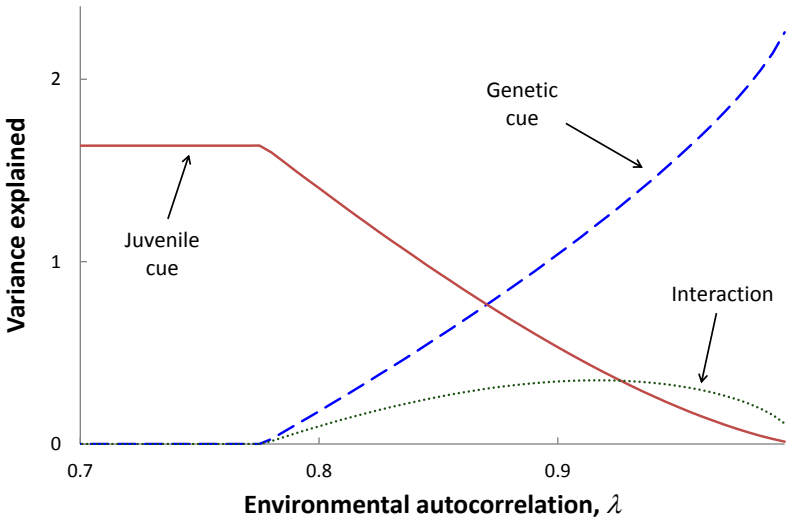


Figure 4a

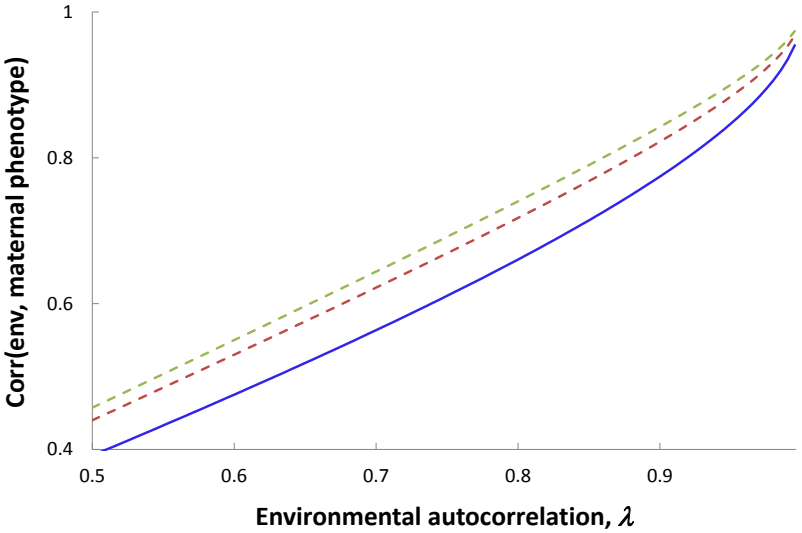


Figure 4b

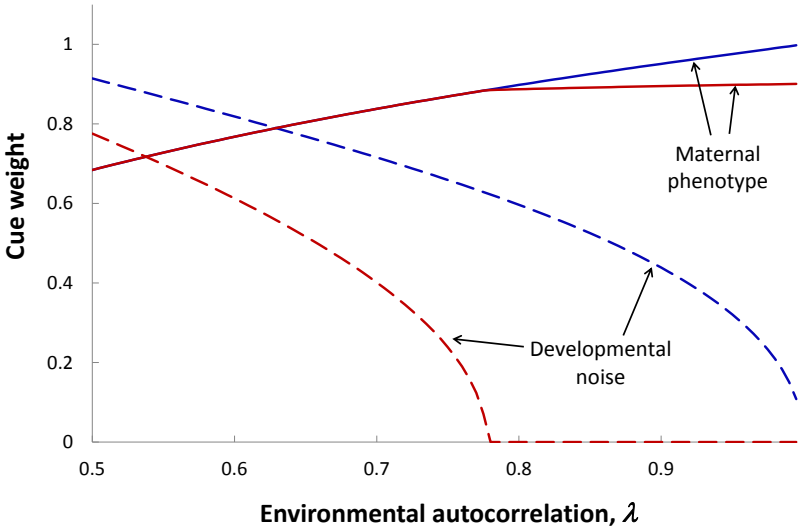


Figure 4c

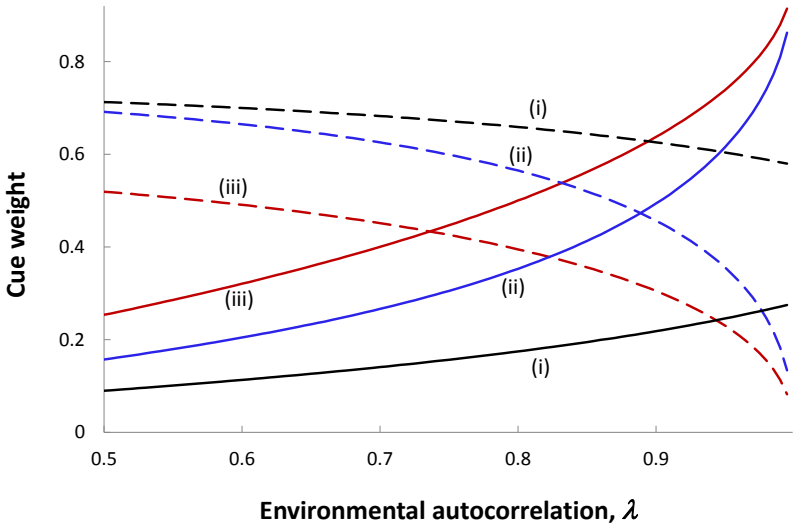

0.8
0.6
0.4
0.2
0
Cue weight
(i)
(ii)
(iii)
(iii)
(ii)
(i)
0.5
0.6
0.7
0.8
0.9
Environmental autocorrelation, λ

Figure 5a

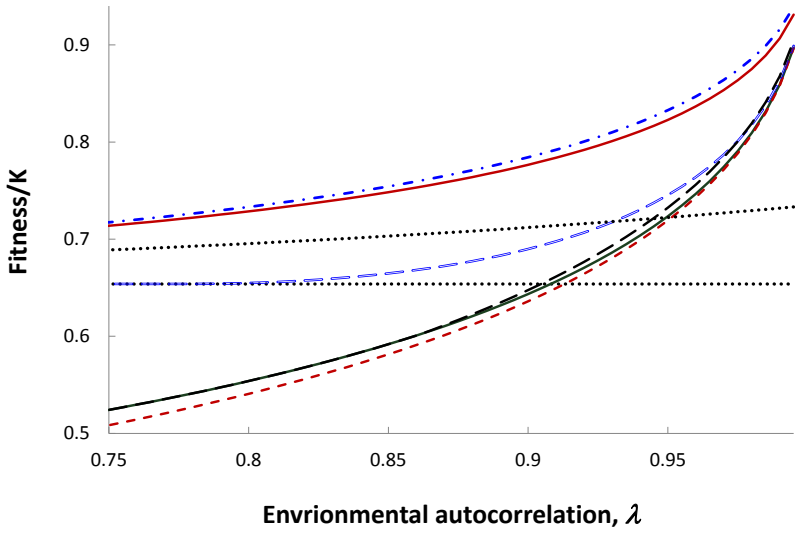
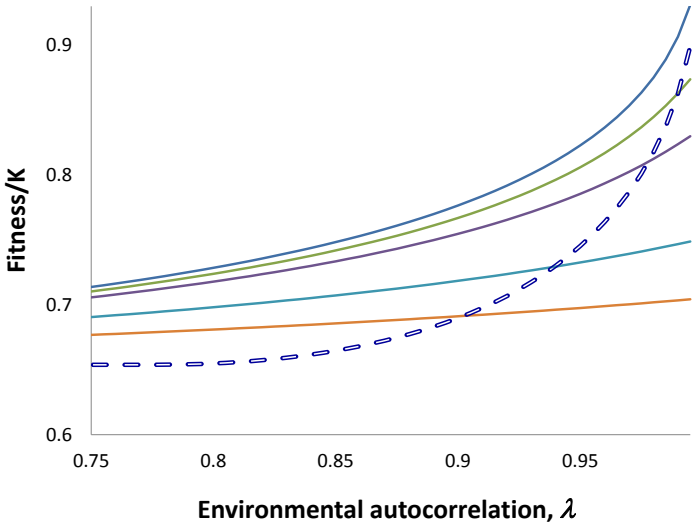


Figure 5b



Supporting Information.

As in the main text we suppose that the adult phenotype of an individual is given by

$$x = \alpha z + \beta_J c_J + \beta_A c_A + \gamma(m + \epsilon_m) + \delta\epsilon_\delta, \tag{11}$$

where z is the value of its genetic effect, c_J is its juvenile environmental cue, c_A is the adult environmental cue observed by her mother, m is the phenotype of the mother, $\epsilon_m \sim N(0, \sigma_m^2)$ is the error in transmission of the maternal phenotype to the offspring and $\epsilon_\delta \sim N(0, 1)$ is a developmental noise term. Here α, β_J, β_A, γ and δ are non-negative genetically determined parameters that specify the action of the developmental system. We analyse the dynamics over time of a large (infinite) cohort of individuals all of which have a developmental system specified by the same parameters $\alpha, \beta_J, \beta_A, \gamma, \delta$.

Consider the characteristics of a randomly chosen member of generation t. The following three random variables are central to our analysis:
$X(t)$ = phenotype of the individual.
$Z(t)$ = genetic cue trait of the individual.
$M(t)$ = phenotype of individual's mother.
We make the assumption that the joint distribution of $X(0)$ and $Z(0)$ is bivariate normal.

SI.1. Change in the joint distribution of X and Z over one generation

In this section we show that this joint distribution remains bivariate normal in subsequent generations. We also derive equations showing how within-generations means and variances change.

We begin by conditioning on a realisation of the stochastic process $\{\theta(t) : t = 0, 1, 2, \ldots\}$.

Lemma 1

Suppose that the joint distribution of $X(t)$ and $Z(t)$ is bivariate normal with:

$\mathbb{E}(X(t)) = \bar{x}$,

$\mathbb{E}(Z(t)) = \bar{z}$,

$\mathrm{Var}(X(t)) = \sigma_X^2$,

$\mathrm{Var}(Z(t)) = \sigma_Z^2$,

$\mathrm{Cov}(X(t), Z(t)) = \sigma_{XZ}$.

Then the joint distribution of $M(t+1)$ and $Z(t+1)$ is bivariate normal with:

$\mathbb{E}(M(t+1)) = \frac{\bar{x}+\theta(t)\sigma_X^2}{1+\sigma_X^2}$,

$\mathbb{E}(Z(t+1)) = \bar{z} + (\frac{\sigma_{XZ}}{1+\sigma_X^2})(\theta(t) - \bar{x})$,

$\mathrm{Var}(M(t+1)) = \frac{\sigma_X^2}{1+\sigma_X^2}$,

$\mathrm{Var}(Z(t+1)) = \sigma_{mut}^2 + \sigma_Z^2 - \frac{\sigma_{XZ}^2}{1+\sigma_X^2}$,

$\mathrm{Cov}(M(t+1), Z(t+1)) = \frac{\sigma_{XZ}}{1+\sigma_X^2}$.

Proof of lemma 1. It is convenient to set $\Delta = \sigma_X^2\sigma_Z^2 - \sigma_{XZ}^2$. Let $f_{XZ}(x, z)$ denote the joint probability density function (pdf) of the two random variables $X(t)$ and $Z(t)$. Then since the joint distribution is bivariate normal we have

$$-2\ln f_{XZ}(x, z) = K_{XX}x^2 - 2K_{XZ}xz + K_{ZZ}z^2 + 2K_X x + 2K_Z z + \text{constant}, \quad (12)$$

where

$K_{XX} = \frac{\sigma_Z^2}{\Delta}$,

$K_{XZ} = \frac{\sigma_{XZ}}{\Delta}$,

$K_{ZZ} = \frac{\sigma_X^2}{\Delta}$,

$K_X = \frac{(\sigma_{XZ}\bar{z} - \sigma_Z^2\bar{x})}{\Delta}$,

$K_Z = \frac{(\sigma_{XZ}\bar{x} - \sigma_X^2 \bar{z})}{\Delta}$.

Consider the distribution of those offspring produced in this generation. Let M be the phenotype of the parent of a randomly selected offspring (this random variable is $M(t+1)$). Also let Z' be the genetic cue value of a randomly selected offspring before mutation of the genetic cue genes, which is the genetic cue value of its mother. Let $f_{MZ'}(m, z')$ denote the joint pdf of M and Z'. Then because of differential number of recruits this density function is proportional to $f_{XZ}(m, z')e^{-(m-\theta)^2/2}$. Thus

$$-2 \ln f_{MZ'}(m, z') = (K_{XX}+1)m^2 - 2K_{XZ}mz' + K_{ZZ}z'^2 + 2(K_X - \theta(t))m + 2K_Z z' + \text{constant}. \tag{13}$$

This is the pdf of a bivariate normal distribution where

$K_{XX} + 1 = \frac{\sigma_{Z'}^2}{\hat{\Delta}}$,

$K_{XZ} = \frac{\sigma_{MZ'}}{\hat{\Delta}}$,

$K_{ZZ} = \frac{\sigma_M^2}{\hat{\Delta}}$,

$K_X - \theta(t) = \frac{(\sigma_{MZ'}\bar{z'} - \sigma_{Z'}^2 \bar{m})}{\hat{\Delta}}$,

$K_Z = \frac{(\sigma_{MZ'}\bar{m} - \sigma_M^2 \bar{z'})}{\hat{\Delta}}$,

and where $\hat{\Delta} = \sigma_M^2 \sigma_{Z'}^2 - \sigma_{MZ'}^2$, $\mathbb{E}(M) = \bar{m}$ and $\mathbb{E}(Z') = \bar{z'}$. From the first three of these equations we deduce that

$$\hat{\Delta} = \frac{\Delta}{1 + \sigma_X^2}. \tag{14}$$

Thus, from these three equations we have

$$\sigma_M^2 = \frac{\sigma_X^2}{1 + \sigma_X^2}, \tag{15}$$

$$\sigma_{MZ'} = \frac{\sigma_{XZ}}{1 + \sigma_X^2}, \tag{16}$$

$$\sigma^2_{Z'} = \sigma^2_Z - \frac{\sigma^2_{XZ}}{1+\sigma^2_X}. \tag{17}$$

From the remaining two equations we then have

$$\bar{m} = \frac{\bar{x} + \theta(t)\sigma^2_X}{1+\sigma^2_X}, \tag{18}$$

$$\bar{z'} = \bar{z} + (\frac{\sigma_{XZ}}{1+\sigma^2_X})(\theta(t) - \bar{x}). \tag{19}$$

Adding mutation to the genetic cue value then gives the results stated in the lemma.

Lemma 2

Suppose that the joint distribution of $X(t)$ and $Z(t)$ is bivariate normal with:

$\mathbb{E}(X(t)) = \bar{x}$,

$\mathbb{E}(Z(t)) = \bar{z}$,

$\mathrm{Var}(X(t)) = \sigma^2_X$,

$\mathrm{Var}(Z(t)) = \sigma^2_Z$,

$\mathrm{Cov}(X(t), Z(t)) = \sigma_{XZ}$.

Then the joint distribution of $X(t+1)$ and $Z(t+1)$ is bivariate normal with:

$$\mathbb{E}(X(t+1)) = \alpha\bar{z} + \beta_J\theta(t+1) + \beta_A\theta(t) + \frac{1}{1+\sigma^2_X}\left[(\alpha\sigma_{XZ} + \gamma\sigma^2_X)\theta(t) + (\gamma - \alpha\sigma_{XZ})\bar{x}\right], \tag{20}$$

$$\mathbb{E}(Z(t+1)) = \bar{z} + (\frac{\sigma_{XZ}}{1+\sigma^2_X})(\theta(t) - \bar{x}), \tag{21}$$

$$\mathrm{Var}(X(t+1)) = \alpha^2(\sigma^2_{mut} + \sigma^2_Z) + (\frac{1}{1+\sigma^2_X})(\gamma^2\sigma^2_X - \alpha^2\sigma^2_{XZ} + 2\alpha\gamma\sigma_{XZ}) + \eta^2, \tag{22}$$

$$\mathrm{Var}(Z(t+1) = \sigma^2_{mut} + \sigma^2_Z - \frac{\sigma^2_{XZ}}{1+\sigma^2_X}, \tag{23}$$

$$\mathrm{Cov}(X(t+1), Z(t+1)) = \alpha(\sigma^2_{mut} + \sigma^2_Z) + (\frac{\sigma_{XZ}}{1+\sigma^2_X})(\gamma - \alpha\sigma_{XZ}), \tag{24}$$

where $\eta^2 = \beta^2_J\sigma^2_J + \beta^2_A\sigma^2_A + \gamma^2\sigma^2_m + \delta^2$.

Proof of lemma 2. Since phenotypes are determined via $x = \alpha z + \beta_J c_J + \beta_A c_A + \gamma(m + \epsilon_m) + \delta\epsilon_\delta$, we see that $X(t+1)$ can be expressed as

$$X(t+1) = \alpha Z(t+1) + \gamma M(t+1) + V, \tag{25}$$

where $V = \beta_J C_J + \beta_A C_A + \gamma\epsilon_m + \delta\epsilon_\delta$. Since $C_J \sim N(\theta(t+1), \sigma_J^2)$ and $C_A \sim N(\theta(t), \sigma_A^2)$ we have $V \sim N(\beta_J\theta(t+1) + \beta_A\theta(t), \eta^2)$. Note that V is conditionally independent of $Z(t+1)$ and $M(t+1)$ given the process $\{\theta(t) : t = 0, 1, 2, \ldots\}$. Thus the joint distribution of $X(t+1)$ and $Z(t+1)$ is bivariate normal.

From this decomposition we have

$$\mathbb{E}(X(t+1) = \alpha\mathbb{E}(Z(t+1)) + \gamma\mathbb{E}(M(t+1)) + \beta_J\theta(t+1) + \beta_A\theta(t). \tag{26}$$

Thus by Lemma 1

$$\mathbb{E}(X(t+1)) = \alpha\bar{z} + \beta_J\theta(t+1) + \beta_A\theta(t) + \frac{1}{1+\sigma_X^2}\left[(\alpha\sigma_{XZ} + \gamma\sigma_X^2)\theta(t) + (\gamma - \alpha\sigma_{XZ})\bar{x}\right]. \tag{27}$$

This establishes equation (20). Equations (21) and (23) were already proved in Lemma 1. To prove equation (22) we note that from equation (25) that we have

$$\mathrm{Var}(X(t+1)) = \alpha^2\mathrm{Var}(Z(t+1)) + \gamma^2\mathrm{Var}(M(t+1)) + 2\alpha\gamma\mathrm{Cov}(Z(t+1), M(t+1)) + \eta^2. \tag{28}$$

The result then follows by substituting the values of $\mathrm{Var}(Z(t+1))$, $\mathrm{Var}(M(t+1))$ and $\mathrm{Cov}((Z(t+1), M(t+1))$ from Lemma 1. From equation (25) we also have

$$\mathrm{Cov}(X(t+1), Z(t+1)) = \alpha\mathrm{Var}(Z(t+1)) + \gamma\mathrm{Cov}(M(t+1), Z(t+1)). \tag{29}$$

Equation (24) then follow by substituting from Lemma 1.

By our assumptions that the joint distribution of $X(0)$ and $Z(0)$ is bivariate normal and Lemma 2, we deduce that the joint distribution of is $X(t)$ and $Z(t)$ also bivariate normal in every generation.

SI.2. Equilibrium variance and covariance values within a generation

The change in variance and covariance values in the above is independent of the environmental process. By Lemma 2 the value at $t+1$ (primed quantities) can be expressed in terms of the value at t as

$$\sigma_X'^2 = \alpha^2(\sigma_{mut}^2 + \sigma_Z^2) + (v^{-2})(\gamma^2\sigma_X^2 - \alpha^2\sigma_{XZ}^2 + 2\alpha\gamma\sigma_{XZ}) + \eta^2, \tag{30}$$

$$\sigma_Z'^2 = \sigma_{mut}^2 + \sigma_Z^2 - v^{-2}\sigma_{XZ}^2, \tag{31}$$

$$\sigma_{XZ}' = \alpha(\sigma_{mut}^2 + \sigma_Z^2) + v^{-2}\sigma_{XZ}(\gamma - \alpha\sigma_{XZ}). \tag{32}$$

where $v^2 = 1 + \sigma_X^2$. To investigate whether these quantities tend to limiting values over time we performed the following calculation. First note that by multipling both sides of equation (31) by α^2 and both sides of equation (32) by α then σ_{mut} only appears in terms where it is a product with α. Thus, without loss of generality we can scale quantities so that $\sigma_{mut} = 1$. We then chose 10000 combinations of the parameters α, γ and η^2, where for each combination the values of these three parameters was chosen independently from a uniform distribution on the interval $(0, 2)$. For each parameter combination we chose 10000 combinations of the initial values of σ_X^2, σ_Z^2 and σ_{XZ}, where for each combination σ_X^2 and σ_Z^2 were both chosen independently from a uniform distribution on the interval $(0, 10)$ and we set $\sigma_{XZ} = r\sigma_X\sigma_Z$, where the correlation coefficient r was chosen independently from a uniform distribution on the interval $(-1, 1)$. For each of these 10^8 combinations of parameters and initial values we iterated the above updating scheme N times, recording the

absolute difference between the final and penultimate values of the three variables σ_X^2, σ_Z^2 and σ_{XZ}. We then set d_{XX} to be the maximum over all 10^8 runs of these absolute differences for σ_X^2, with d_{XZ} and d_{ZZ} similarly defined. When $N = 1000$ the values of d_{XX}, d_{XZ} and d_{ZZ} were all less than 10^{-10}. However, when we set $N = 100$ we noted that for those combinations of parameter values with small α the covariance term was slow to converge, presumably because there is then weak selection on the quantitative genetic effect. We therefore repeated our similation with the values of α chosen independently from a uniform distribution on the interval $(0.1, 2)$. For this simulation we recorded $d_{XX} = 0.00000068$, $d_{XZ} = 0.00001133$ and $d_{ZZ} = 0.00000175$.

Given the above simulations it seems reasonable to assume that the iterative scheme for σ_X^2, σ_Z^2 and σ_{XZ} converges, albeit rather slowly for small values of α. We therefore make this assumption and seek the limiting values of these quantities analytically. To do so we set $\sigma_X'^2 = \sigma_X^2$, $\sigma_Z'^2 = \sigma_Z^2$ and $\sigma_{XZ}' = \sigma_{XZ}$ to obtain three simultaneous equations. From equation (31) we obtain

$$\sigma_{XZ} = v\sigma_{mut}. \tag{33}$$

Feeding this into equations (30) and (32) we obtain

$$v^4 - \alpha\sigma_{mut}v^3 - (\gamma^2 + 1 + \eta^2)v^2 - \alpha\gamma\sigma_{mut}v + \gamma^2 = 0. \tag{34}$$

Since $v^2 = 1 + \sigma_X^2$, we seek a solution of this equation in the range $v > 1$.

Lemma 3. (Existence and uniqueness)
Equation (34) has a unique solution in the range $v > 1$.

Proof of lemma 3. To investigate whether there exist a solution to equation

(34) we set

$$f(v) = v^4 - \alpha\sigma_{mut}v^3 - (\gamma^2 + 1 + \eta^2)v^2 - \alpha\gamma\sigma_{mut}v + \gamma^2. \tag{35}$$

Then it is easily verified that $f(1) < 0$ and that $f(v) \to \infty$ as $v \to \infty$. Thus there must exist $v > 1$ such that $f(v) = 0$. We denote the minimum such value by $\hat{v}$. Note that $f(v) < 0$ for $1 \leq v < \hat{v}$ and $f(\hat{v}) = 0$. It follows that $f'(\hat{v}) \geq 0$.

We next show that this $\hat{v}$ cannot be a double root. To do so we note that

$$vf''(v) - 3f'(v) = 3\alpha\sigma_{mut}v^2 + 4(\gamma^2 + 1 + \eta^2) + 3\alpha\gamma\sigma_{mut}. \tag{36}$$

Since all the coefficients on the right hand side of this equation are non-negative, and at least one is positive we have $\hat{v}f''(\hat{v}) > 3f'(\hat{v})$. Thus, if we had $f'(\hat{v}) = 0$, then this would imply that $f''(\hat{v}) > 0$, so that f would have a strict local minimum at $v = \hat{v}$, contradicting the fact that $f(v) < 0$ for $1 \leq v < \hat{v}$. It follows that we must have $f'(\hat{v}) > 0$.

Set $h(v) = vf'(v) - 4f(v)$. Note that since $f(\hat{v}) = 0$ and $f'(\hat{v}) > 0$ we have $h(\hat{v}) > 0$. Now suppose that there is at least one further root of equation (34) that is greater than $\hat{v}$. Let v_1 be the minimum such root. Then since $f'(\hat{v}) > 0$ we must have $f(v) > 0$ for $\hat{v} < v < v_1$. Since $f(v_1) = 0$ we thus have $f'(v_1) \leq 0$. Its follows that $h(v_1) \leq 0$. But from the definition of the function f we have

$$h(v) = \alpha\sigma_{mut}v^3 + 2(\gamma^2 + 1 + \eta^2)v^2 + 3\alpha\gamma\sigma_{mut}v - 4\gamma^2. \tag{37}$$

so that $h(v)$ is a strictly increasing function of v. Thus contradict the fact that $h(\hat{v}) > 0$ and $h(v_1) \leq 0$. We conclude that there is no such root v_1 and that the equation $f(v) = 0$ has a unique solution for $v \geq 1$.

SI.3. The vector process

Set:

$\bar{X}(t)$ = mean phenotype in generation t.

$\bar{Z}(t)$ = mean genetic cue value in generation t.

Here we derive the dynamics of the vector stochastic process $\{(\theta(t), \bar{X}(t), \bar{Z}(t)) : t = 0, 1, 2, \ldots\}$.

From equation (33) we can write

$$\mathbb{E}(X(t+1)) = \alpha\bar{Z}(t)+\beta_J\theta(t+1)+(\beta_A+\gamma)\theta(t)+(\gamma v^{-2}-\alpha v^{-1}\sigma_{mut})(\bar{X}(t)-\theta(t)) \quad (38)$$

and

$$\mathbb{E}(Z(t+1)) = \bar{Z}(t) + v^{-1}\sigma_{mut}(\theta(t) - \bar{X}(t)). \quad (39)$$

From the above we see that the vector process $\{(\theta(t), \bar{X}(t), \bar{Z}(t)) : t = 0, 1, 2, \ldots\}$ has dynamic equations given by

$$\theta(t+1) = \lambda\theta(t) + \epsilon_\theta(t), \quad (40)$$

$$\bar{X}(t+1) = \alpha\bar{Z}(t) + (\lambda + A)\theta(t) + (B - L)(\bar{X}(t) - \theta(t)) + \beta_J\epsilon_\theta(t) \quad (41)$$

$$\alpha\bar{Z}(t+1) = \alpha\bar{Z}(t) + L(\theta(t) - \bar{X}(t)). \quad (42)$$

where

$A = \lambda(\beta_J - 1) + \beta_A + \gamma$,

$B = v^{-2}\gamma$,

$L = \alpha v^{-1}\sigma_{mut}$.

SI.4. Equilibrium across-generational variances and covariances for mean values

We now assume that vector stochastic process $\{(\theta(t), \bar{X}(t), \bar{Z}(t)) : t = 0, 1, 2, \ldots\}$ has a stationary distribution, and find the various variances and covariances of the components at this equilibrium.

Set

$D(t) = \bar{X}(t) - \theta(t)$,

$\hat{Z}(t) = \alpha \bar{Z}(t)$,

and consider the vector process $\{(\theta(t), D(t), \hat{Z}(t)) : t = 0, 1, 2, \ldots\}$. By the equations for the process $\{(\theta(t), \bar{X}(t), \bar{Z}(t)) : t = 0, 1, 2, \ldots\}$ we have

$$\theta(t+1) = \lambda\theta(t) + \epsilon_\theta(t), \tag{43}$$

$$D(t+1) = \hat{Z}(t) + A\theta(t) + (B - L)D(t) + (\beta_J - 1)\epsilon_\theta(t), \tag{44}$$

$$\hat{Z}(t+1) = \hat{Z}(t) - LD(t). \tag{45}$$

We now assume stationarity in these equations so that means and variances do not depend on t. From equation (45) we see that

$$\mathbb{E}(\hat{Z}(t+1)) = \mathbb{E}(\hat{Z}(t)) - L\mathbb{E}(D(t)). \tag{46}$$

Thus assuming stationarity, so that $\mathbb{E}(\hat{Z}(t+1)) = \mathbb{E}(\hat{Z}(t))$ we have $\mathbb{E}(D) = 0$. Thus

$$\mathbb{E}(\bar{X}) = \mathbb{E}(\theta) = 0. \tag{47}$$

From equation (43) we have

$$\mathrm{Var}(\theta(t+1)) = \lambda^2 \mathrm{Var}(\theta(t)) + \sigma^2. \tag{48}$$

Thus assuming that $\mathrm{Var}(\theta(t+1)) = \mathrm{Var}(\theta(t)) = \mathrm{Var}(\theta)$ we have

$$\mathrm{Var}(\theta) = \frac{\sigma^2}{1-\lambda^2}; \tag{49}$$

i.e. equation (2) of the main text. Taking variances in equation (45) we similarly have

$$\mathrm{Var}(\hat{Z}) = \mathrm{Var}(\hat{Z}) - 2L\mathrm{Cov}(\hat{Z}, D) + L^2\mathrm{Var}(D), \tag{50}$$

and hence

$$\mathrm{Cov}(\hat{Z}, D) = \frac{L}{2}\mathrm{Var}(D). \tag{51}$$

From the equations (43) and (45) we have

$$\mathrm{Cov}(\theta, \hat{Z}) = \lambda \mathrm{Cov}(\theta, \hat{Z}) - \lambda L \mathrm{Cov}(\theta, D), \tag{52}$$

and hence

$$\mathrm{Cov}(\theta, \hat{Z}) = -\frac{\lambda L}{1-\lambda}\mathrm{Cov}(\theta, D). \tag{53}$$

From the equations (43) and (44) we have

$$\mathrm{Cov}(\theta, D) = \lambda \mathrm{Cov}(\theta, \hat{Z}) + \lambda A \mathrm{Var}(\theta) + \lambda(B-L)\mathrm{Cov}(\theta, D) + (\beta_J - 1)\sigma^2. \tag{54}$$

After rearranging and substituting for $\mathrm{Cov}(\theta, Z)$ in terms of $\mathrm{Cov}(\theta, D)$ from equation (53), and for $\mathrm{Var}(\theta)$ from equation (49) we get

$$\mathrm{Cov}(\theta, D) = \left(\frac{\sigma^2}{1+\lambda}\right) \frac{\lambda(\beta_A + \gamma) + \beta_J - 1}{1 - \lambda + (\lambda^2 - \lambda)B + \lambda L}, \tag{55}$$

where we have made use of the definition of A. From equations (44) and (45) we have

$$\begin{aligned}\mathrm{Cov}(\hat{Z},D) &= \mathrm{Var}(\hat{Z}) + A\mathrm{Cov}(\theta,\hat{Z}) + (B-L)\mathrm{Cov}(\hat{Z},D) \\ &- L\mathrm{Cov}(\hat{Z},D) - AL\mathrm{Cov}(\theta,D) - (B-L)L\mathrm{Var}(D). \end{aligned} \quad (56)$$

After some rearrangement and substitutions from equations (51) and (53) we get

$$\mathrm{Var}(\hat{Z}) = \frac{L}{2}(1+B)\mathrm{Var}(D) + \frac{AL}{1-\lambda}\mathrm{Cov}(\theta,D). \quad (57)$$

Finally, taking variances on both sides of equation (44), rearranging and substituting we have

$$\begin{aligned}\left[1 - \frac{L}{2}(1-B) - B^2\right]\mathrm{Var}(D) &= \left(\frac{A^2\sigma^2}{1-\lambda^2}\right) + (1-\beta_J)^2\sigma^2 \\ &+ \left(\frac{A}{1-\lambda}\right)[2B(1-\lambda) - L]\,\mathrm{Cov}(\theta,D). \end{aligned} \quad (58)$$

Thus, $\mathrm{Var}(D)$ can be found from this equation and equation (55). Since $D(t) = \bar{X}(t) - \theta(t)$ it is then possible to find the fitness of the developmental system from equation (9) of the main text. For later convenience we define $g(\alpha,\beta_J,\beta_A,\gamma,\delta) = \ln G(\alpha,\beta_J,\beta_A,\gamma,\delta)$ and express g as

$$g(\alpha,\beta_J,\beta_A,\gamma,\delta) = \ln K - \frac{1}{2}\ln(v^2) - \frac{\mathrm{Var}(D)}{2v^2}, \quad (59)$$

where $v^2 = 1 + \sigma_X^2$.

SI.5. Special case: optimal randomisation when there are no cues

Suppose that there are no cues, i.e $\alpha = \beta_J = \beta_A = \gamma = 0$ so that phenotype

determination is given by $x = \delta\epsilon_\delta$. We analyse this special case, investigating the optimal amount of randomisation.

In this case we have $A = -\lambda$, $B = 0$ and $L = 0$. Thus by equation (58) we have $\mathrm{Var}(D) = \mathrm{Var}(\theta)$. From equation (34) we also have $v^2 = 1 + \eta^2$, so that $v^2 = 1 + \delta^2$. Let $\hat{g}(\delta) = g(0, 0, 0, 0, \delta)$ be the logarithm of geometric mean fitness for randomisation δ. Then from equation (59) we have

$$\hat{g}(\delta) = \ln K - \frac{1}{2}\ln(v^2) - \frac{\mathrm{Var}(\theta)}{2v^2}. \tag{60}$$

Differentiating we have

$$\hat{g}'(\delta) = \frac{\delta}{v^4}[\mathrm{Var}(\theta) - (1 + \delta^2)]. \tag{61}$$

Thus the optimal value of δ is $\delta^* = 0$ when $\mathrm{Var}(\theta) < 1$ and $\delta^* = \sqrt{\mathrm{Var}(\theta) - 1}$ when $\mathrm{Var}(\theta) \geq 1$.

SI.6. Special case: the juvenile cue only and the need for randomisation

Suppose that there is just the juvenile environmental cue; i.e. $\alpha = \beta_A = \gamma = 0$, but there may be randomisation, so that phenotype determination is given by $x = \beta_J c_J + \delta\epsilon_\delta$. We analyse this special case, deriving inequalities for the value of β_J^* and then use this inequality to show that $\delta^* = 0$.

In this case we have $A = \lambda(\beta_J - 1)$, $B = 0$ and $L = 0$. Thus by equation (58) we have $\mathrm{Var}(D) = (1 - \beta_J)^2\mathrm{Var}(\theta)$. From equation (34) we also have $v^2 = 1 + \eta^2$, so that $v^2 = 1 + \beta_J\sigma_J^2 + \delta^2$. Let $\hat{g}(\beta_J, \delta) = g(0, \beta_J, 0, 0, \delta)$ be the logarithm of geometric

mean fitness. By equation (59)

$$\hat{g}(\beta_J, \delta) = \ln K - \frac{1}{2}\ln(v^2) - \frac{(1-\beta_J)^2 \mathrm{Var}(\theta)}{2v^2}. \tag{62}$$

Differentiating we have

$$v^4 \frac{\partial \hat{g}}{\partial \beta_J} = \mathrm{Var}(\theta)[(\beta_J - 1)^2 \beta_J \sigma_J^2 - (\beta_J - 1)v^2] - \beta_J \sigma_J^2 v^2. \tag{63}$$

It follows that $\frac{\partial \hat{g}}{\partial \beta_J}(0, \delta) = v^{-2} > 0$ for all δ. Thus $\beta_J^* > 0$.

Since the optimal value of β_J^* is positive we have $\frac{\partial \hat{g}}{\partial \beta_J}(\beta_J^*, \delta^*) = 0$. Thus from equation (63) we have

$$\mathrm{Var}(\theta)[(\beta_J^* - 1)^2 \beta_J^* \sigma_J^2 - (\beta_J^* - 1)v^{*2}] = \beta_J^* \sigma_J^2 v^{*2}, \tag{64}$$

where $v^{*2} = 1 + (\beta_J^*)^2 + (\delta^*)^2$. We thus have $(\beta_J^* - 1)^2 \beta_J^* \sigma_J^2 > (\beta_J^* - 1)v^{*2}$, so that

$$(\beta_J^* - 1)^2 (\beta_J^*)^2 \sigma_J^2 > \beta_J^* (\beta_J^* - 1)v^{*2}. \tag{65}$$

It follows that $\beta_J^* \neq 1$. Furthermore, since $v^{*2} > \beta_J^{*2} \sigma_J^2$ we have

$$(\beta_J^* - 1)^2 > \beta_J^* (\beta_J^* - 1), \tag{66}$$

from which it easily follows that since $\beta_J \neq 1$ then $\beta_J^* < 1$. Overall we conclude that $0 < \beta_J^* < 1$.

We now focus on δ^*. From equation (62) we have

$$v^4 \frac{\partial \hat{g}}{\partial \delta}(\beta_J^*, \delta^*) = \delta^*[(\beta_J^* - 1)^2 \mathrm{Var}(\theta) - v^{*2}]. \tag{67}$$

However, from equation (64) we have

$$\beta_J^* \sigma_J^2 [\mathrm{Var}(\theta)(\beta_J^* - 1)^2 - v^{*2}] = (\beta_J^* - 1) v^{*2} \mathrm{Var}(\theta) < 0, \tag{68}$$

so that $\mathrm{Var}(\theta)(\beta_J^* - 1)^2 - v^{*2} < 0$. It follows that $\frac{\partial \hat{g}}{\partial \delta}(\beta_J^*, \delta^*) < 0$, which implies that $\delta^* = 0$.

SI.7. Special case: maternal adult cue only

Suppose that there is just the maternal adult environmental cue; i.e. $\alpha = \beta_J = \gamma = 0$, but there may be randomisation, so that phenotype determination is given by $x = \beta_A c_A + \delta \epsilon_\delta$. We analyse this special case, deriving inequalities for the value of β_A^*.

In this case we have $A = \beta_A - \lambda$, $B = 0$ and $L = 0$. Thus by equation (58) we have $\mathrm{Var}(D) = (\beta_A^2 + 1 - 2\lambda\beta_A)\mathrm{Var}(\theta)$. From equation (34) we also have $v^2 = 1 + \eta^2$, so that $v^2 = 1 + \beta_A \sigma_A^2 + \delta^2$. Let $\hat{g}(\beta_A, \delta) = g(0, 0, \beta_A, 0, \delta)$ be the logarithm of geometric mean fitness. By equation (59)

$$\hat{g}(\beta_A, \delta) = \ln K - \frac{1}{2}\ln(v^2) - \frac{(\beta_A^2 + 1 - 2\lambda\beta_A)\mathrm{Var}(\theta)}{2v^2}. \tag{69}$$

Differentiating we have

$$v^4 \frac{\partial \hat{g}}{\partial \beta_A} = \mathrm{Var}(\theta)[(\beta_A^2 + 1 - 2\lambda\beta_A)\beta_A \sigma_A^2 - (\beta_A - \lambda)v^2] - \beta_A \sigma_A^2 v^2. \tag{70}$$

It follows that $\frac{\partial \hat{g}}{\partial \beta_A}(0, \delta) = \lambda v^2 \mathrm{Var}(\theta) > 0$ for all δ. Thus $\beta_A^* > 0$.

Since the optimal value of β_A^* is positive we have $\frac{\partial \hat{g}}{\partial \beta_A}(\beta_A^*, \delta^*) = 0$. Thus from equation (70) we have

$$\mathrm{Var}(\theta)[(\beta_A^{*2} + 1 - 2\lambda\beta_A^*)\beta_A^* \sigma_A^2 - (\beta_A^* - \lambda)v^{*2}] = \beta_A^* \sigma_A^2 v^{*2}. \tag{71}$$

where $v^{*2} = 1 + (\beta_A^*)^2\sigma_A^2 + (\delta^*)^2$. We thus have $(\beta_A^{*2} + 1 - 2\lambda\beta_A^*)\beta_A^*\sigma_A^2 > (\beta_A^* - \lambda)v^{*2}$, so that

$$(\beta_A^{*2} + 1 - 2\lambda\beta_A^*)\beta_A^{*2}\sigma_A^2 > \beta_A^*(\beta_A^* - \lambda)v^{*2}. \tag{72}$$

Now suppose that $\lambda\beta_A^* \geq 1$. Then since this implies we have $\beta_A^* > \lambda$ we have $(\beta_A^* - \lambda)v^{*2} > (\beta_A^* - \lambda)\beta_A^{*2}\sigma_A^2$, so that $(\beta_A^{*2} + 1 - 2\lambda\beta_A^*) > \beta_A^*(\beta_A^* - \lambda)$. It follows that $\lambda\beta_A^* < 1$. This proves that $\lambda\beta_A^* < 1$. Overall we conclude that $0 < \beta_A^* < \frac{1}{\lambda}$.

SI.8. Special case: some inequalities on the optimal juvenile and maternal adult cue weights

Suppose that there are just two cues; the juvenile cue and the adult maternal cue; i.e. $\alpha = \gamma = 0$, but there may be randomisation, so that phenotype determination is given by $x = \beta_J c_J + \beta_A c_A + \delta\epsilon_\delta$. We analyse this special case, deriving inequalities for the value of β_J^* and β_A^*. We first show that these weights are positive.

In this case we have $A = \lambda(\beta_J - 1) + \beta_A$, $B = 0$ and $L = 0$. Thus by equation (58) we have $\mathrm{Var}(D) = H\mathrm{Var}(\theta)$ where

$$H = (\beta_J - 1)^2 + \beta_A^2 + 2\lambda\beta_A(\beta_J - 1). \tag{73}$$

Let $\hat{g}(\beta_J, \beta_A, \delta) = g(0, \beta_J, \beta_A, 0, \delta)$ be the logarithm of geometric mean fitness. By equation (59)

$$\hat{g}(\beta_J, \beta_A, \delta) = \ln K - \frac{1}{2}\ln(v^2) - \frac{H\mathrm{Var}(\theta)}{2v^2}. \tag{74}$$

Differentiating we have

$$v^4\frac{\partial\hat{g}}{\partial\beta_J} = \mathrm{Var}(\theta)[H\beta_J\sigma_J^2 - (\beta_J - 1 + \lambda\beta_A)v^2] - \beta_J\sigma_J^2v^2. \tag{75}$$

Thus

$$v^4 \frac{\partial \hat{g}}{\partial \beta_J}(0, \beta_A^*, \delta^*) = \mathrm{Var}(\theta)[(1 - \lambda\beta_A^*)v^2]. \tag{76}$$

Now suppose that $\beta_J^* = 0$. Then the analysis of Section SI.7 shows that $\lambda\beta_A^* < 1$. Thus $\frac{\partial \hat{g}}{\partial \beta_J}(0, \beta_A^*, \delta^*) > 0$, contradicting the fact that $\beta_J^* = 0$. We deduce that $\beta_J^* > 0$. Similarly

$$v^4 \frac{\partial \hat{g}}{\partial \beta_A} = \mathrm{Var}(\theta)[H\beta_A\sigma_A^2 - (\beta_A + \lambda(\beta_J - 1))v^2] - \beta_A\sigma_A^2 v^2. \tag{77}$$

Thus

$$v^4 \frac{\partial \hat{g}}{\partial \beta_A}(\beta_J^*, 0, \delta^*) = \mathrm{Var}(\theta)[(\lambda(1 - \beta_J^*)v^2], \tag{78}$$

and a similar argument using the results of SI.6 shows that $\beta_A^* > 0$.

From the above we may assume that $\frac{\partial g}{\partial \beta_J}(\beta_J^*, \beta_A^*, \delta^*) = 0$ and $\frac{\partial \hat{g}}{\partial \beta_A}(\beta_J^*, \beta_A^*, \delta^*) = 0$. Thus

$$\mathrm{Var}(\theta)[H^*\beta_J^*\sigma_J^2 - (\beta_J^* - 1 + \lambda\beta_A^*)v^{*2}] = \beta_J^*\sigma_J^2 v^{*2} \tag{79}$$

and

$$\mathrm{Var}(\theta)[H^*\beta_A^*\sigma_A^2 - (\beta_A^* + \lambda(\beta_J^* - 1))v^{*2}] = \beta_A^*\sigma_A^2 v^{*2}, \tag{80}$$

where

$$H^* = (\beta_J^* - 1)^2 + \beta_A^{*2} + 2\lambda\beta_A^*(\beta_J^* - 1) \tag{81}$$

and $v^{*2} = 1 + \beta_J^{*2}\sigma_J^2 + \beta_A^{*2}\sigma_A^2$. Multiplying each side of equation (79) by β_J^*, both sides of equation (80) by β_A^* and adding the two resulting equations gives $\mathrm{Var}(\theta)M = (\beta_J^{*2}\sigma_J^2 + \beta_A^{*2}\sigma_A^2)v^{*2}$, where

$$M = H^*[\beta_J^{*2}\sigma_J^2 + \beta_A^{*2}\sigma_A^2] - [\beta_J^*(\beta_J^* - 1 + \lambda\beta_A^*) + \beta_A^*(\beta_A^* + \lambda(\beta_J^* - 1))]v^{*2}. \tag{82}$$

Note that $M = [\beta_J^{*2}\sigma_J^2 + \beta_A^{*2}\sigma_A^2]v^{*2}/\mathrm{Var}(\theta) > 0$. Thus the term on the right hand

side of equation (82) is positive. Since we can write H^* as $H^* = (\beta_J^* + \lambda\beta_A^* - 1)^2 + (1-\lambda^2)\beta_A^{*2}$, we have $H^* > 0$. Since $\beta_J^{*2}\sigma_J^2 + \beta_A^{*2}\sigma_A^2 < v^{*2}$ we deduce that

$$H^* > \beta_J^*(\beta_J^* - 1 + \lambda\beta_A^*) + \beta_A^*(\beta_A^* + \lambda(\beta_J^* - 1)). \tag{83}$$

After expanding terms this yields

$$\beta_J^* + \lambda\beta_A^* < 1. \tag{84}$$

We now return to equations (79) and (80). These equations can be written as

$$\beta_J^* - 1 + \lambda\beta_A^* = \beta_J^*\sigma_J^2\left[\frac{H*}{v^{*2}} - \frac{1}{\mathrm{Var}(\theta)}\right] \tag{85}$$

and

$$\lambda(\beta_J^* - 1) + \beta_A^* = \beta_A^*\sigma_A^2\left[\frac{H*}{v^{*2}} - \frac{1}{\mathrm{Var}(\theta)}\right]. \tag{86}$$

Thus $\beta_J^* - 1 + \lambda\beta_A^*$ and $\lambda(\beta_J^* - 1) + \beta_A^*$ have the same sign. From equation (84) we deduce that

$$\lambda\beta_J^* + \beta_A^* < \lambda. \tag{87}$$

SI.9. The maternal phenotype is always used as a cue when the other cues are the juvenile and maternal adult cue

Suppose that there are three cues; the juvenile cue, the adult maternal cue and the maternal phenotype as a cue; i.e. $\alpha = 0$, and may be randomisation and transmission error, so that phenotype determination is given by $x = \beta_J c_J + \beta_A c_A + \gamma(m + \epsilon_m) + \delta\epsilon_\delta$. Here we show that there is always positive weight assigned to the maternal phenotype as a cue under optimal phenotype detwermination; i.e. $\gamma^* > 0$. To show this let β_J^*, β_A^* and δ^* be the optimal weights when γ is constrained to be zero. Set

$\hat{g}(\gamma) = g(0, \beta_J^*, \beta_A^*, \gamma, \delta^*)$. Then we will show that $\hat{g}'(0) > 0$.

From equation (59) we can write

$$\hat{g}(\gamma) = \ln K - \frac{1}{2}\ln(v^2(\gamma)) - \frac{\mathrm{Var}(D)(\gamma)}{2v^2(\gamma)}, \tag{88}$$

where we now regard $v(\gamma)$ and $\mathrm{Var}(D)(\gamma)$ as functions of γ. From equation (34) we have $v(0) = 1 + \beta_J^{*2}\sigma_J^2 + \beta_A^{*2}\sigma_A^2 + \delta^{*2}$. By implicit differentiation of equation (34) with respect to γ it can also be verified that $v'(0) = 0$. By equation (58) we have $\mathrm{Var}(D)(0) = H^*\mathrm{Var}(\theta)$ where H^* is given by equation (81). By implicit differentiation of equation (58) with respect to γ it can also be verified that

$$\mathrm{Var}(D)'(0) = 2(\lambda(\beta_J^* - 1) + \beta_A^*)\left[\mathrm{Var}(\theta) + \frac{1}{v^2(0)}\mathrm{Cov}(\theta, D)(0)\right]. \tag{89}$$

By equation (55) we also have $\mathrm{Cov}(\theta, D) = \mathrm{Var}(\theta)(\beta_J^* + \lambda\beta_A^* - 1)$. Thus

$$\mathrm{Var}(D)'(0) = 2(\lambda(\beta_J^* - 1) + \beta_A^*)\mathrm{Var}(\theta)\left[1 + \frac{1}{v^2(0)}(\beta_J^* + \lambda\beta_A^* - 1)\right]. \tag{90}$$

From equation (88) we then have

$$v^4(0)\hat{g}'(0) = -(\lambda(\beta_J^* - 1) + \beta_A^*)\mathrm{Var}(\theta)\left[v^2(0) + (\beta_J^* + \lambda\beta_A^* - 1)\right]. \tag{91}$$

Note that since $v > 1$ we have $v^2(0) + (\beta_J^* + \lambda\beta_A^* - 1) > 0$. Also by inequality (87) we have $\lambda(\beta_J^* - 1) + \beta_A^* < 0$. Thus $\hat{g}'(0) > 0$. It follows that $\gamma^* > 0$.

We note that a similar calculation shows that when the juvenile environmental cue and maternal phenotype are the only cues, we also have $\gamma^* > 0$.